



109950469 1120701

JC05 Rec'd PCT/PTO 03 DEC 2001

PCT

FORM PTO-1390 (modified) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE
(REV 10-95)ATTORNEY'S DOCKET NUMBER
01/22924TRANSMITTAL LETTER TO THE UNITED STATES
DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING
A FILING UNDER 35 U.S.C., 371

U.S. APPLICATION NO. (IF KNOWN), SEE 37 CFR

09/980469

INTERNATIONAL APPLICATION NO.
PCT/IL00/00330INTERNATIONAL FILING DATE
7 June, 2000PRIORITY DATE CLAIMED
10 June, 1999

TITLE OF INVENTION

**PROCESS OF EXPRESSING AND ISOLATING RECOMBINANT PROTEINS AND
RECOMBINANT PROTEIN PRODUCTS FROM PLANTS, PLANT DERIVED TISSUES OR
CULTURED PLANT CELLS**

APPLICANT(S) FOR DO/EO/US

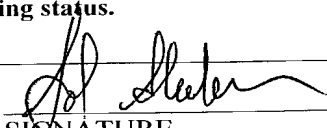
1) SHANI, Ziv 2) SHOSEYOV, Oded

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371
3. ☒ This is an express request to begin national examination procedures (35 U.S.C. 371(f) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(3)(2)
 - a. ☒ is transmitted herewith (required only if not transmitted by the International Bureau).
 - b. ☒ has been transmitted by the International Bureau.
 - c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ A translation of the International Application into English (35 U.S.C. 371(3)(2).
7. ☒ A copy of the International Search Report (PCT/ISA/210).
8. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3)
 - a. ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
 - b. ☐ have been transmitted by the International Bureau.
 - c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
 - d. ☐ have not been made and will not be made.
9. ☐ A translation of the amendments to the claims under PCT Article 19(35 U.S.C. 371(c)(3)).
10. ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
11. ☒ A copy of the International Preliminary Examination Report (PCT/IPEA/409).
12. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 13 to 18 below concern document(s) of information included:

13. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
14. ☒ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
15. ☐ A FIRST preliminary amendment.
A SECOND or SUBSEQUENT preliminary amendment.
16. ☐ A substitute specification.
17. ☐ A change of power of attorney and/or address letter.
18. ☐ Certificate of Mailing by Express Mail
19. ☐ Sequence Listing Statement: The sequence listing information recorded in computer readable form is identical to the written (on paper or compact disc) sequence listing.
20. ☐ Other items or information:

U.S. APPLICATION NO. (IF KNOWN) Sec 37 CFR 097 980469		INTERNATIONAL APPLICATION NO. PCT/IL00/00330		ATTORNEY'S DOCKET NUMBER 01/22924	
20. The following fees are submitted: BASIC NATIONAL FEE (37 CFR 1.492(a) (1) – (5)): <input type="checkbox"/> Search Report has been prepared by the EPO or JPO \$ 890 <input type="checkbox"/> International preliminary examination fee paid to USPTO (37 CFR 1,482) \$ 710 <input type="checkbox"/> No international preliminary examination fee paid to USPTO (37 CFR 1,482) but international search fee paid to USPTO (37 CFR 1,445(a)(2)) \$ 740 <input type="checkbox"/> Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO \$ 1040 <input checked="" type="checkbox"/> International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4) \$ 100 ENTER APPROPRIATE BASIC FEE AMOUNT =				CALCULATIONS PTO USE ONLY <div style="border: 1px solid black; padding: 2px;"> \$100.00 </div>	
Surcharge of \$130.00 for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(e))				<div style="border: 1px solid black; padding: 2px;"> \$0.00 </div>	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total claims	25 - 20 =	5	5 x \$ 18	\$90.00	
Independent claims	6 - 3 =	3	3 x \$ 84	\$252.00	
Multiple Dependent Claims (check if applicable) <input type="checkbox"/>				\$0.00	
TOTAL OF ABOVE CALCULATIONS =				\$442.00	
Reduction of 1/2 for filing by small entity, if applicable. <input checked="" type="checkbox"/>				\$221.00	
SUBTOTAL =				\$221.00	
Processing fee of \$130.00 for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492(f))				<div style="border: 1px solid black; padding: 2px;"> \$0.00 </div>	
TOTAL NATIONAL FEE =				\$221.00	
Fee for recording the enclosed assignment (37 CFR 1.2(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31) (check if applicable) <input checked="" type="checkbox"/>				<div style="border: 1px solid black; padding: 2px;"> \$40.00 </div>	
TOTAL FEES ENCLOSED =				\$261.00	
				Amount to be refunded: \$	
				charged \$	
<input type="checkbox"/> A check in the amount of \$ _____ to cover the above fees is enclosed. <input checked="" type="checkbox"/> Please charge my Deposit Account No. 50-1407 in the amount of \$ 261.00 to cover the above fees. A duplicate copy of this sheet is enclosed. <input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge any fees which may be required, or credit any overpayment to Deposit Account No. 50-1407 . A duplicate copy of this sheet is enclosed. NOTE: Where an appropriate time limit under 37CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status. SEND ALL CORRESPONDENCE TO:					
SOL SHEINBEIN G.E. EHRLICH (1995) LTD. C/O ANTONHY CASTORINA SUITE 207 2001 JEFFERSON DAVIS HIGHWAY ARLINGTON, VIRGINIA 22202, USA			<div style="text-align: center;">  _____ SIGNATURE </div> <div style="text-align: center;"> SOL SHEINBEIN _____ NAME </div> <div style="text-align: center;"> 25,457 _____ REGISTRATION NUMBER </div> <div style="text-align: center;"> Nov 13, 2001 _____ DATE </div>		

1

PROCESS OF EXPRESSING AND ISOLATING RECOMBINANT
PROTEINS AND RECOMBINANT PROTEIN PRODUCTS FROM
PLANTS, PLANT DERIVED TISSUES OR CULTURED PLANT CELLS

5 FIELD AND BACKGROUND OF THE INVENTION

The present invention relates to a process of expressing and isolating recombinant proteins and recombinant protein products from plants, plant derived tissue or cultured plant cells, which process exploits (i) the high affinity between cellulose binding peptides and cellulose; (ii) the inherent
10 abundance of cellulose in *planta*; and/or (iii) the simplicity associated with cellulose isolation from plants, plant derived tissue and cultured plant cells.

More particularly, the present invention relates to a process expressing and isolating recombinant proteins and recombinant protein products from plants, plant derived tissue or cultured plant cells, which
15 process employs the expression of a fusion protein including a recombinant protein and a cellulose binding peptide fused thereto, plant homogenization, isolation of a fusion protein cellulosic matter complex and optional subsequent isolation of the fusion protein and/or the recombinant protein from the complex. The present invention further relates to nucleic acid
20 molecules and to genetically modified or viral infected plants or plant cells which are useful while implementing the process, and further to a novel composition of matter which results from the process.

Citation or identification of any reference in this section or in any other section of this application shall not be construed as an admission that
25 such reference is available as prior art to the present invention.

With the advent of recombinant technology, techniques for the genetic transformation of various host organisms, such as bacteria, yeasts, fungi, plants and animals, for the purposes of producing specific proteins through the expression of heterologous or foreign genes have been
30 extensively developed.

Using these recombinant techniques and hosts, numerous commercially important recombinant proteins (examples of which are included hereinbelow) have been expressed and purified. Expression and isolation of a protein of interest on a commercial scale, necessitate the
35 selection of a suitable expression host. This selection largely depends on the economics of production and purification, as well as the ability of the host to accomplish the post-translational modifications needed for full biological activity of the recombinant protein.

Much of the early work in biotechnology was directed toward the expression of recombinant or "heterologous" proteins in prokaryotes like *Escherichia coli* and *Bacillus subtilis* because of the relative ease of genetic manipulation, growth in batch culture and large-scale fermentation of prokaryotes.

Although *E. coli* can in certain cases perform some post translational modifications and events, such as, protein folding and disulfide bond formation, it cannot secrete proteins extracellularly nor can it glycosylate, gamma carboxylate, beta hydroxylate, acetylate or process pre- and pro-peptides. *B. subtilis* suffers from the same limitations as *E. coli* except that it is capable of extracellular secretion.

Furthermore, *E. coli* and other bacteria are pathogens and therefore, depending on the application, contaminants such as pyrogens and endotoxins expressed along with the recombinant protein must be removed. In addition, extensive post-purification chemical and enzymatic treatments (e.g., to refold the protein into an active form) are sometimes required in order to obtain a biologically active protein.

Because proteins are not secreted from prokaryotes like *E. coli*, bacterial cells must be disrupted for product recovery. The subsequent release of bacterial contaminants and other proteins make product purification more difficult and expensive. Because purification accounts for up to 90 % of the total cost of producing recombinant proteins in bacteria, proteins like Insulin can cost several thousand dollars per gram when recombinantly produced in, and subsequently isolated from, *E. coli*.

Because of the many limitations associated with prokaryotic hosts, the biotechnology industry has looked for eukaryotic host cultures such as, yeast, fungi, insect cells, and mammalian cell tissue culture, to properly and efficiently express recombinant proteins.

For most of the proteins requiring extensive post-translational modifications for therapeutic and/or functional activity, mammalian cell culture is the most common alternative to *E. coli*. Although mammalian cells are capable of correctly folding and glycosylating bioactive proteins, the quality and extent of glycosylation can vary with different culture conditions among the same host cells. Furthermore, mammalian culture has extremely high fermentation costs (60-80% of total production expense), requires expensive media, and poses safety concerns from potential contamination by viruses and other pathogens. Yields are generally low and

in the range of 0.5-1.5% of cellular protein, or micrograms per liter (up to 300-400 milligrams per liter).

Yeast, other fungi, and insect cells are currently being used as alternatives to mammalian cell culture. Yeast, however, produces
5 incorrectly glycosylated proteins that have excessive mannose residues and are generally limited in eukaryotic processing. Further, although the baculovirus insect cell system can produce high levels of glycosylated proteins, these are typically not secreted, making purification complex and expensive. Fungi represent the best current system for high-volume, low-
10 cost production of recombinant proteins, but they are not capable of expressing many target proteins.

In addition, eukaryotic cultures, require the maintenance of suitable conditions for efficient commercially viable expression of proteins. As such, the ambient temperature, pH value and aeration level of such cultures
15 need to be carefully controlled, while nutrients must be added to the culture medium in carefully regulated doses and waste products removed. In addition, rigorous aseptic practices must be observed in order to avoid contamination by extraneous microbes. Such cultures are thus normally grown in sophisticated fermentors or bioreactors which are housed in
20 expensively maintained factories. Such overheads are reflected in the high price of the recombinant protein end-products.

To a lesser extent, animals have also been utilized for the production of recombinant proteins. Although large amounts of protein can be produced and relatively easily recovered from such animals (e.g., proteins
25 specifically produced in mammary glands and secreted with the milk), production in such host is limited to the expression of proteins which do not interfere with the host physiology. In addition, transgenic animals are subject to lengthy lead times to develop herds with stable genetics, high operating costs, contamination by animal viruses and a relatively slow rate
30 of biomass generation substantially prolonging the time period following which recovery of commercial amounts of the protein can be effected.

The biochemical, technical and economic limitations on existing prokaryotic and eukaryotic expression systems has created substantial interest in developing new expression systems for the production of
35 recombinant proteins.

Plants represent the most likely alternative to existing expression systems. With the availability and on going development of plant

transformation techniques, most commercially important plant species can now be genetically modified to express a variety of recombinant proteins.

Such transformation techniques include, for example, the *Agrobacterium* vector system, which involves infection of the plant tissue with a bacterium (*Agrobacterium*) into which the foreign gene has been inserted. A number of methods for transforming plant cells with *Agrobacterium* are well known (Klee *et al.*, Annu. Rev. Plant Physiol. (1987) 38:467-486; Schell and Vasil Academic Publishers, San Diego, Calif. (1989) p. 2-25; and Gatenby (1989) in Plant Biotechnology, eds. Kung, S. and Arntzen, C. J., Butterworth Publishers, Boston, Mass. p. 93-112).

The biolistic or particle gun method, which permits genetic material to be delivered directly into intact cells or tissues by bombarding regeneratable tissues, such as meristems or embryogenic callus, with DNA-coated microparticles has contributed to plant transformation simplicity and efficiency. The microparticles penetrate the plant cells and act as inert carriers of a genetic material to be introduced therein. Microprojectile bombardment of embryogenic suspension cultures has proven successful for the production of transgenic plants of a variety of species. Various parameters that influence DNA delivery by particle bombardment have been defined (Klein *et al.*, Bio/Technology (1998) 6:559-563; McCabe *et al.*, Bio/Technology (1998) 6:923-926; and Sanford, Physiol. Plant. (1990) 79:206-209).

Micropipette systems are also used for the delivery of foreign DNA into plants via microinjection (Neuhaus *et al.*, Theor. Appl. Genet. (1987) 75:30-36; and Neuhaus and Spangenberg, Physiol. Plant. (1990) 79:213-217).

Other techniques developed to introduce foreign genes into plants include direct DNA uptake by plant tissue, or plant cell protoplasts (Schell and Vasil (1987) Academic Publishers, San Diego, Calif. p. 52-68; and Toriyama *et al.*, Bio/Technology (1988) 6:1072-1074) or by germinating pollen (Chapman, Mantell and Daniels (1985) W. Longman, London, p. 197-209; and Ohta, Proc. Natl. Acad. Sci. USA (1986) 83:715-719).

DNA uptake induced by brief electric shock of plant cells has also been described (Zhang *et al.*, Plant. Cell. Rep. (1988) 7:379-384 and Fromm *et al.*, Nature (1986) 319:791-793).

In addition, virus mediated plant transformation has also been extensively described. Transformation of plants using plant viruses is

described, for example, in U.S. Pat. No. 4,855,237 (BGV), EP-A 67,553 (TMV), Japanese Published Application No. 63-14693, EPA 194,809, EPA 278,667, and Gluzman *et al.*, (1988) Communications in Molecular Biology: Viral Vectors, Cold Spring Harbor Laboratory, New York, pp. 172-189. Pseudovirus particles for use in expressing foreign DNA in many hosts, including plants, have also been described, for reference, see, for example WO 87/06261.

The production of recombinant proteins and peptides in plants has been investigated using a variety of approaches including transcriptional fusions using a strong constitutive plant promoter (e.g., from cauliflower mosaic virus, Sijmons *et al.*, Bio/Technology (1990) 8:217-221); transcriptional fusions with organ specific promoter sequences (Radke *et al.*, Theoret. Appl. Genet. (1988) 75:685-694); and translational fusions which require subsequent cleavage of a recombinant protein (Vanderkerckove *et al.*, Bio/Technology (1989) 7:929-932).

The application of such genetic transformation techniques has allowed the incorporation of a variety of important genetic traits for crop improvement and also for the biotechnological production of extractable, valuable, foreign proteins including enzymes, vaccine proteins and antibodies.

Foreign proteins that have been successfully expressed in plant cells include proteins from bacteria (Fraley *et al.* Proc. Natl. Acad. Sci. U.S.A. (1993) 80:4803-4807), animals (Misra and Gedamu, Theor. Appl. Genet. (1989) 78:161-168), fungi and other plant species (Fraley *et al.* Proc. Natl. Acad. Sci. U.S.A. (1983) 80:4803-4807). Some proteins, predominantly markers of DNA integration, have been expressed in specific cells and tissues including seeds (Sen Gupta-Gopalan *et al.* Proc. Natl. Acad. Sci. U.S.A. (1985) 82:3320-3324; Radke *et al.* Theor. Appl. Genet. (1988) 75:685-694).

Due to the advantageous economics of field-grown crops, the ability to synthesize proteins in storage organs like tubers, seeds, fruits and leaves and the ability of plants to perform many of the post-translational modifications previously described, several plant expression systems are currently investigated for potential as highly effective and economically feasible systems for the production of recombinant proteins.

Since highly expressive systems such as the ubiquitin fusion system described in U.S. Pat. No. 5,773,705 have been demonstrated, a major hurdle to an effective plant expression system resides with the relatively

As such, alternative expression approaches have been undertaken in an effort to simplify the purification procedure of the recombinant protein from the plant cells.

Another system utilizing seeds as an expression host is disclosed in U.S. Pat. No. 5,888,789. This system provides for the secretion of heterologous protein by malting of monocot plant seeds. The heterologous genes are expressed during germination of the seeds and isolated from a malt.

U.S. Pat No. 5,650,554 describes the use of a class of genes called oil body protein genes, that have unique features, allowing the production of recombinant proteins that can be easily separated from other host cell components.

Although systems which target or direct recombinant protein production to specific tissues allow for easier recombinant protein isolation such systems are typically limited in the effective host range and/or the amounts of recombinant proteins produced since such systems fail to exploit the entire plant biomass.

A novel approach for simplifying the purification of recombinant enzymes from plant host cells is disclosed in U.S. Pat. No. 5,474,925 which describes an expression construct utilizing a signal peptide translationally fused to a recombinant enzyme which targets the enzyme to the cellulose matrix of the cell wall. This enables the isolation of the

enzyme along with the easily recoverable cellulose matrix. This system is utilized for the localized expression of commercially important enzymes in cotton fibers. According to this system, the expressed enzymes are recovered along with the cellulosic matter of the fibers. The enzyme-cellulose matrix recovered, is directly utilized for commercial enzymatic processes.

Although this system presents a simple means with which a recombinant protein can be expressed and isolated, it is limited to the production of enzymes in cotton fibers of the cotton plant.

10 Furthermore, a major hurdle encountered when expressing cellulose
targeted proteins within a plant is the interference of the expressed products
in the natural formation of the cell wall, which typically results in growth
arrest of the plant growth. Although this hurdle can be overcome by, for
example, targeting the protein to specific plant tissue as is the case for U.S.
15 Pat. No. 5,474,925, this targeting severely limits the expressing biomass
and as such the quantity of the expressed protein. In addition, targeting the
expression to a specific plant tissue also limits the number of plant species
which can be effectively utilized for such an expression.

There is thus a widely recognized need for, and it would be highly
20 advantageous to have, a plant expression system and method which provide
high level of expression of a recombinant protein and which allow simple
and effective recovery of the expressed recombinant protein devoid of the
above limitations.

25 SUMMARY OF THE INVENTION

According to one aspect of the present invention there is provided a process of expressing a recombinant protein in a plant and of isolating the recombinant protein from the plant, the process comprising the steps of (a) providing a plant, a plant derived tissue or cultured plant cells expressing a fusion protein including the recombinant protein and a cellulose binding peptide being fused thereto, the fusion protein being compartmentalized within cells of the plant, plant derived tissue or cultured plant cells, so as to be sequestered from cell walls of the cells of the plant, plant derived tissue or cultured plant cells; (b) homogenizing the plant, plant derived tissue or cultured plant cells, so as to bring into contact the fusion protein with a cellulosic matter of the plant, plant derived tissue or cultured plant cells, to thereby effect affinity binding of the fusion protein via the cellulose binding peptide to the cellulosic matter, thereby obtaining a fusion protein cellulosic

matter complex; and (c) isolating the fusion protein cellulosic matter complex.

According to further features in preferred embodiments of the invention described below, the process further comprising the steps of
5 washing the fusion protein cellulosic matter complex, thereby removing endogenous plant proteins and other plant material therefrom and collecting the fusion protein cellulosic matter complex as a final product of the process.

According to still further features in the described preferred
10 embodiments the process further comprising the steps of washing the fusion protein cellulosic matter complex, thereby removing endogenous plant proteins and other plant material therefrom, exposing the fusion protein cellulosic matter complex to conditions effective in dissociating the fusion protein from the cellulosic matter; and isolating the fusion protein, thereby
15 obtaining an isolated fusion protein.

According to still further features in the described preferred embodiments the process further comprising the steps of exposing the isolated fusion protein to conditions effective in digesting the fusion protein so as to release the recombinant protein therefrom, thereby obtaining a
20 released recombinant protein and isolating the released recombinant protein.

According to still further features in the described preferred embodiments the process further comprising the steps of washing the fusion protein cellulosic matter complex, thereby removing endogenous plant proteins and other plant material therefrom, exposing the fusion protein
25 cellulosic matter complex to conditions effective in digesting the fusion protein so as to release the recombinant protein therefrom, thereby obtaining a released recombinant protein, and isolating the released recombinant protein.

According to still further features in the described preferred
30 embodiments, the conditions effective in dissociating the fusion protein from the cellulosic matter are selected from the group consisting of basic conditions, denaturative conditions and affinity displacement conditions.

According to still further features in the described preferred embodiments, the conditions effective in digesting the fusion protein so as
35 to release the recombinant protein therefrom are selected from the group consisting of proteolysis effected via a protease and proteolysis effected under predetermined cis or trans conditions for digesting a controllable intervening protein sequence.

According to another aspect of the present invention there is provided a genetically modified or viral infected plant or cultured plant cells expressing a fusion protein including a recombinant protein and a cellulose binding peptide

5 According to further features in preferred embodiments of the invention described below, the fusion protein is compartmentalized within cells of the plant or cultured plant cells, so as to be sequestered from cell walls of the cells of the plant or cultured plant cells.

According to still further features in the described preferred
10 embodiments the fusion protein is compartmentalized within a cellular
compartment selected from the group consisting of cytoplasm, endoplasmic
reticulum, golgi apparatus, oil bodies, starch bodies, chloroplastids,
chloroplasts, chromoplastids, chromoplasts, vacuole, lysosomes,
mitochondria, and nucleus.

15 According to still further features in the described preferred
embodiments expression of the fusion protein is under a control of a
constitutive or tissue specific plant promoter.

According to still further features in the described preferred
embodiments the fusion protein includes a recombinant protein and a
cellulose binding peptide separated therebetween via a unique amino acid
sequence recognizable and digestible by a protease or under predetermined
cis or trans conditions for digesting a controllable intervening protein
sequence.

According to yet another aspect of the present invention there is provided a composition of matter comprising (a) a plant derived cellulosic matter; and (b) a fusion protein including a recombinant protein and a cellulose binding peptide separated therebetween via a unique amino acid sequence recognizable and digestible by a protease or under predetermined cis or trans conditions for digesting a controllable intervening protein sequence, the fusion protein being complexed to the plant derived cellulosic matter by affinity binding via the cellulose binding peptide.

According to still another aspect of the present invention there is provided a nucleic acid molecule comprising (a) a promoter sequence for directing protein expression in plant cells; (b) a heterologous nucleic acid sequence including (i) a first sequence encoding a cellulose binding peptide; (ii) a second sequence encoding a recombinant protein, wherein the first and second sequences are joined together in frame; optionally (iii) a third sequence encoding a signal peptide for directing a protein to a cellular

compartment, the third sequence being upstream and in frame with the first and second sequences; and/or optionally (iv) a fourth sequence encoding a unique amino acid sequence being recognizable and digestible by a protease or under predetermined cis or trans conditions for digesting a controllable intervening protein sequence, the fourth sequence being between and in frame with the first and second sequences, wherein, the heterologous nucleic acid sequence being down stream the promoter sequence, such that expression of the heterologous nucleic acid sequence is effectable by the promoter sequence.

According to further features in preferred embodiments of the invention described below, the nucleic acid molecule further comprising a sequence element selected from the group consisting of an origin of replication for propagation in bacterial cells, at least one sequence element for integration into a plant's genome, a polyadenylation recognition sequence, a transcription termination signal, a sequence encoding a translation start site, a sequence encoding a translation stop site, plant RNA virus derived sequences, plant DNA virus derived sequences, tumor inducing (Ti) plasmid derived sequences, and a transposable element derived sequence.

The present invention successfully addresses the shortcomings of the presently known configurations by providing a novel process of expressing and isolating recombinant proteins and recombinant protein products from plants, plant derived tissue or cultured plant cells, which process exploits (i) the high affinity between cellulose binding peptides and cellulose; (ii) the inherent abundance of cellulose in *planta*; and (iii) the simplicity associated with cellulose isolation from plants, plant derived tissue and/or cultured plant cells.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention is herein described, by way of example only, with reference to the accompanying drawings. With specific reference now to the drawings in detail, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of the preferred embodiments of the present invention only, and are presented in the cause of providing what is believed to be the most useful and readily understood description of the principles and conceptual aspects of the invention. In this regard, no attempt is made to show structural details of the invention in more detail than is necessary for a fundamental understanding of the

invention, the description taken with the drawings making apparent to those skilled in the art how the several forms of the invention may be embodied in practice.

In the drawings:

5 FIG. 1 is a process flow chart demonstrating the process according to the present invention.

FIGs. 2a-b are nucleotide (2a) and restriction maps (2b) of the insert into the pUC19-cell1-ProtL-cexNG vector constructed in accordance with the teachings of the present invention.

10 FIGs. 3a-b are nucleotide (3a) and restriction maps (3b) of the insert into the pBS-Sig-cex-Fx-HDEL vector constructed in accordance with the teachings of the present invention.

FIGs. 4a-b are nucleotide (4a) and restriction maps (4b) of the insert into the pBS-Sig-Tma-Fx-HDEL vector constructed in accordance with the teachings of the present invention.

15 FIG. 5 is a flow chart diagram depicting the step involved in the analysis of the transgenic plant material produced according to the teachings of the present invention.

FIG. 6 is a gel image of PCR amplified fragments from ProtL-cex transformants 1-15. N - negative control, M - molecular weight marker (MWM).

FIGs. 7a-b are gel images of PCR amplified fragments from cex-fx transformants 1-19. N - negative control, M - molecular weight marker (MWM).

25 FIGs. 8a-b are gel images of PCR amplified fragments from Tma-fx transformants 1-19. C - positive control, M - molecular weight marker (MWM).

FIGs. 9a-b are immunoblot images of proteins extracted from ProtL-cexNG transformants. The extracted proteins were separated on SDS-PAGE, blotted, and reacted with anti CBDcex antibody(6a) or mouse IgG (6b) as is further described in Example 2. M - MWM, a -a WT plant cell wall fraction, b - transformant line 2 cell wall fraction, c - WT cellulose fraction (exogenous), d - transformant line 2 cellulose fraction.

35 FIGs. 10a-b are immunoblot images of proteins extracted from CBDcex-fx transformants. The extracted proteins were separated on SDS-PAGE, blotted, and reacted with anti-Fx (7a) or anti-CBDcex (7b) antibodies as is further described in Example 2. C - positive control, M - MWM, wt - WT plant line, 5-24 - transformant plant lines.

FIGs. 11a-d are immunoblot images of proteins extracted from CDBTma-fx transformants. The extracted proteins were separated on SDS-PAGE, blotted, and reacted with anti-Fx antibodies as is further described in Example 2. Figures 11a and 11c represent proteins extracted from the cellulose fraction, while Figures 11b and 11d represent proteins extracted from the wall fraction. C - positive control, M - MWM, wt - WT plant line, 2-19 - transformant plant lines.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is of a process which can be used for expressing and isolating recombinant proteins and recombinant protein products from plants, plant derived tissue or cultured plant cells. The present invention is further of nucleic acid molecules and genetically modified or viral infected plants or plant cells which are useful while implementing the process, and further of a novel composition of matter which results from the process. Specifically, the present invention can be used to obtain large quantities of the recombinant proteins and the recombinant protein products in a simple and cost effective manner, since the process according to the present invention exploits (i) the high affinity between cellulose binding peptides and cellulose; (ii) the inherent abundance of cellulose in *planta*; and (iii) the simplicity associated with cellulose isolation from plants, plant derived tissue and/or cultured plant cells.

The principles and operation of a process according to the present invention may be better understood with reference to the drawings and accompanying descriptions.

Before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not limited in its application to the details of construction and the arrangement of the components set forth in the following description or illustrated in the drawings. The invention is capable of other embodiments or of being practiced or carried out in various ways. Also, it is to be understood that the phraseology and terminology employed herein is for the purpose of description and should not be regarded as limiting.

Process:

Referring now to the drawings, Figure 1 is a flow chart illustrating the process according to the teachings of the present invention.

The process according to the present invention is directed at expressing a recombinant protein in a plant and further at isolating the recombinant protein from the plant. The process according to the present invention is effected by first providing a plant, a plant derived tissue or
5 cultured plant cells (which are referred to herein below individually and collectively as "plant material") 10 expressing a fusion protein which includes the recombinant protein and a cellulose binding peptide fused thereto. The fusion protein is compartmentalized within cells of the plant material, so as to be sequestered from cell walls of the cells of the plant
10 material. As used herein in the specification and in the claims section that follows, the phrase "cultured plant cells" includes both non-differentiated plant cell cultures and some what more differentiated callus cultures.

Compartmentalizing and thereby sequestering the fusion protein from the cell walls is an essential feature of the present invention because
15 high levels of expressed cellulose binding peptide associated with plant cell walls inhibit plant growth to a great extent. See to this effect U.S. Pat. applications Nos. 09/006,632; 09/006,636; and PCT/IL98/00345 (WO 99/07830).

When sufficient expression has been detected by sampling and
20 testing the plant material as further detailed hereinunder, the plant material is homogenized 12 so as to bring into contact the fusion protein with a cellulosic matter of the plant material, to thereby effect affinity binding 14 of the fusion protein via the cellulose binding peptide to the cellulosic matter, thereby obtaining a fusion protein cellulosic matter complex.
25 Conditions such as, but not limited to, temperature, pH, salt concentration, time and the like are preferably set so as to allow maximal binding. Such conditions are well known to the skilled artisan and can be experimentally modified to best suit a specific application. Sampling and testing can be employed to monitor the binding process, as further detailed hereinunder.

When sufficient binding has occurred the fusion protein cellulosic
30 matter complex is collected or isolated 16 by methods well known to the skilled artisan which methods are traditionally employed for isolation of cellulosic matter from plant material. Thereafter, a wash step 18 is employed to remove unbound material, including, in particular, unbound
35 endogenous plant proteins, thereby isolating the fusion protein cellulosic matter complex. The wash step can be repeated one or several times with a single or several wash solutions, each of which can include in addition to water, buffers, salts, detergents and the like to efficiently effect the removal

of unbound matter from the fusion protein cellulosic matter complex. The wash step can be effected in solution using appropriate stirring, however, advantageously, the wash step is effected within a column into which the collected or isolated fusion protein cellulosic matter complex is packed and subsequently washed.

According to one embodiment of the present invention, and as indicated in Figure 1 by numeral 20, the fusion protein cellulosic matter complex is collected as a final product of the process. Such a final product can serve as a pack for affinity columns. In this case the recombinant protein is selected to have affinity to a ligand, which can then be affinity purified via a column packed with the fusion protein cellulosic matter complex, in a manner otherwise similar to that described in U.S. Pat. No. 5,474,925, which is incorporated herein by reference. One of the advantages of the process described herein over the teachings of U.S. Pat. No. 5,474,925 is that by sequestering the fusion protein from the cell walls one can achieve very high expression of the fusion protein as compared to the low expression levels practically enabled by U.S. Pat. No. 5,474,925, because no deleterious effect on plant growth is exerted. As a result, the specific activity of the fusion protein cellulosic matter complex formed according to the present invention, i.e., the number of fusion protein molecules per weight of cellulosic matter, is far superior. Further details relating to the effect of high cellulose binding peptide expression on plant development see also PCT/IL98/000345.

According to another embodiment of the present invention, as indicated by numeral 22, the final product of the process according to the present invention is the fusion protein itself 24. Thus, according to this embodiment of the present invention, conditions effective in dissociating the fusion protein from the cellulosic matter are used to effect such dissociation. The dissociated fusion protein is thereafter readily isolated by any conventional separation technique, such as, but not limited to, elution or size separation, such as differential filtration or centrifugation, thereby obtaining an isolated fusion protein. Conditions effective in dissociating the fusion protein from the cellulosic matter include, but are not limited to, basic conditions (e.g., 20 mM Tris pH 12) which are known to dissociate all cellulose binding peptides from cellulose, denaturative conditions, or affinity displacement conditions, e.g., using 200 nM glucose or cellobiose which are known in their ability to elute family IX cellulose binding domains (CBDs). Alternatively, a protein cleavage site can be inserted in the

cellulose binding peptide to facilitate the dissociation of the fusion protein by specific proteolysis, for example. See to this effect and to other uses of CBD-fusion proteins U.S. Pat. Nos. 5,719,044; 5,670,623; 5,856,021; 5,137,819; 5,202,247; 5,340,731; and 5,474,925; and U.S. Pat. applications
5 Nos. 08/788,621; and 08/788,622; EP 0 381 719 B1, and EP application No. 93907724.4. See also the teachings of U.S. Pat. No. 5,834,247, which is further described hereinunder.

As indicated by numeral 26, the fusion protein thus isolated can be exposed to conditions effective in digesting the fusion protein so as to
10 release the recombinant protein therefrom, thereby obtaining a released recombinant protein which can be thereafter isolated as a final product 28. Conditions effective in digesting the fusion protein so as to release the recombinant protein therefrom include, but are not limited to, proteolysis effected via a protease, such as, but not limited to, Factor Xa, enterokinase,
15 thrombin, trypsin, papain, pepsin, chemotrypsin and the like, or proteolysis effected via controllable intervening protein sequence (CIVPS) inserted into or adjacent the cellulose binding peptide, the CIVPS are capable of excision from or cleavage of the peptide under predetermined conditions in cis or in trans, e.g., increase in temperature, exposure to light, unblocking of amino
20 acid residues by dephosphorylation and treatment with chemical reagents or deglycosylation, examples include proteolysis effected under acidic conditions (HCl, e.g., to cleave between Asp and Pro) and proteolysis effected by a proteolysing reagent, such as CNBr to cleave downstream of Met, all as known in the art and/or as further described in U.S. Pat. No.
25 5,834,247, which is incorporated herein by reference.

Isolating final product 28 from other proteolytic products derived, for example, from the cellulose binding peptide, can be effected by any one of a number of protein isolation techniques well known to the skilled artisan, including, but not limited to, affinity separation via, for example, antibodies
30 bound to a solid support, size and/or charge based separation via gel electrophoresis or chromatography, and the like. Additional methods include, but are not limited to, fractionation, gel-filtration, ion-exchange, hydrophobic, and affinity chromatography, ultrafiltration and crystallization.

According to an alternative embodiment of the process of the present
35 invention, as indicated in Figure 1 by numeral 30, the washed fusion protein cellulose matter complex resulting from step 18 is exposed to conditions effective in digesting the fusion protein so as to release the recombinant protein therefrom, thereby obtaining a released recombinant protein. These

conditions are similar to those described with respect to step 26. The released recombinant protein is thereafter readily isolated by any conventional separation technique, such as, but not limited to, displacement or size separation, such as differential filtration or centrifugation, thereby obtaining an isolated recombinant protein final product 32.

Cellulose binding peptides:

As used herein in the specification and in the claims section below, the phrase "cellulose binding peptide" includes peptides e.g., proteins and domains (portions) thereof, which are capable of, when expressed in plant cells, affinity binding to a plant derived cellulosic matter following homogenization and cell rupture. The phrase thus includes, for example, peptides which were screened for their cellulose binding activity out of a library, such as a peptide library or a DNA library (e.g., a cDNA expression library or a display library) and the genes encoding such peptides isolated and are expressible in plants. Yet, the phrase also includes peptides designed and engineered to be capable of binding to cellulose and/or units thereof.

Such peptides include amino acid sequences expressible in plants that are originally derived from a cellulose binding region of, e.g., a cellulose binding protein (CBP) or a cellulose binding domain (CBD). The cellulose binding peptide according to the present invention can include any amino acid sequence expressible in plants which binds to a cellulose polymer. For example, the cellulose binding domain or protein can be derived from a cellulase, a binding domain of a cellulose binding protein or a protein screened for, and isolated from, a peptide library, or a protein designed and engineered to be capable of binding to cellulose or to saccharide units thereof, and which is expressible in plants. The cellulose binding domain or protein can be naturally occurring or synthetic, as long as it is expressible in plants. Suitable polysaccharidases from which a cellulose binding domain or protein expressible in plants may be obtained include β -1,4-glucanases. In a preferred embodiment, a cellulose binding domain or protein from a cellulase or scaffoldin is used. Typically, the amino acid sequence of the cellulose binding peptide expressed in plants according to the present invention is essentially lacking in the hydrolytic activity of a polysaccharidase (e.g., cellulase, chitinase), but retains the cellulose binding activity. The amino acid sequence preferably has less than about 10 % of the hydrolytic activity of the native polysaccharidase; more preferably less than about 5 %, and most preferably less than about 1 % of

the hydrolytic activity of the native polysaccharidase, ideally no activity altogether.

The cellulose binding domain or protein can be obtained from a variety of sources, including enzymes and other proteins which bind to cellulose which find use in the subject invention.

In Table 4 below are listed those binding domains which bind to one or more soluble/insoluble polysaccharides including all binding domains with affinity for soluble glucans (α , β , and/or mixed linkages). The N1 cellulose-binding domain from endoglucanase CenC of *C. fimi* is the only protein known to bind soluble cellosaccharides and one of a small set of proteins which are known to bind any soluble polysaccharides. Also, listed in Tables 1 to 3 are examples of proteins containing putative β -1,3-glucan-binding domains (Table 1); proteins containing Streptococcal glucan-binding repeats (Cpl superfamily) (Table 2); and enzymes with chitin-binding domains, which may also bind cellulose (Table 3). The genes encoding each one of the peptides listed in Tables 1-4 are either isolated or can be isolated as further detailed hereinunder, and therefore, such peptides are expressible in plants. Scaffoldin proteins or portions thereof, which include a cellulose binding domain, such as that produced by *Clostridium cellulovorans* (Shoseyov *et al.*, PCT/US94/04132) can also be used as the cellulose binding peptide expressible in plants according to the present invention. Several fungi, including *Trichoderma* species and others, also produce polysaccharidases from which polysaccharide binding domains or proteins expressible in plants can be isolated. Additional examples can be found in, for example, Microbial Hydrolysis of Polysaccharides, R. A. J. Warren, Annu. Rev. Microbiol. 1996, 50:183-212; and "Advances in Microbial Physiology" R. K. Poole, Ed., 1995, Academic Press Limited, both are incorporated by reference as if fully set forth herein.

Table 1
Overview of proteins containing putative β -1,3 glucan-binding domains

Source (strain)	Protein	accession No.	Ref ¹
Type I			
<i>B. circulans</i> (WL-12)	GLCA1	P23903/M34503/JQ0420	1
<i>B. circulans</i> (IAM 1165)	BglH	JN0772/D17519/S67033	2

Type II

5	<i>Actinomadura</i> sp. (FC7)	XynII	U08894	3
	<i>Arthrobacter</i> sp. (YCWD3)	GLCI	D23668	9
	<i>O. xanthineolytica</i>	GLC	P22222/M60826/A39094	4
	<i>R. faecitabidus</i> (YLM-50)	RP I	Q05308/A45053/D10753	5a,b
	<i>R. communis</i>	Ricin	A12892	6
	<i>S. lividans</i> (1326)	XlnA	P26514/M64551/JS07986	7
10	<i>T. tridentatus</i>	FactorGa	D16622	8

B. : *Bacillus*, O. : *Oerskovia*, R. *faecitabidus* : *Rarobacter faecitabidus*, R. *communis*: *Ricinus communis*, S. : *Streptomyces*, T. : *Tachypleus* (Horseshoe Crab)

15 ¹References:

- 1) Yahata *et al.* (1990) *Gene* **86**, 113-117
- 2) Yamamoto *et al.* (1993) *Biosci. Biotechnol. Biochem.* **57**, 1518-1525
- 3) Harpin *et al.* (1994) *EMBL Data Library*
- 20 4) Shen *et al.* (1991) *J. Biol. Chem.* **266**, 1058-1063
- 5a) Shimoi *et al.* (1992) *J. Biol. Chem.* **267**, 25189-25195
- 5b) Shimoi *et al.* (1992) *J. Biochem* **110**, 608-613
- 6) Horn *et al.* (1989) Patent A12892
- 7) Shareck *et al.* (1991) *Gene* **107**, 75-82
- 25 8) Seki *et al.* (1994) *J. Biol. Chem.* **269**, 1370-1374
- 9) Watanabe *et al.* (1993) *EMBL Data Library*

Table 2
Overview of proteins containing Streptococcal glucan-binding repeats
(Cpl superfamily)

	Source	Protein	Accession No.	Ref. ²
35	<i>S. downei (sobrinus)</i> (0MZ176)	GTF-I	D13858	1
	<i>S. downei (sobrinus)</i> (MFe28)	GTF-I	P11001/M17391	2
	<i>S. downei (sobrinus)</i> (MFe28)	GTF-S	P29336/M30943/A41483	3
	<i>S. downei (sobrinus)</i> (6715)	GTF-I	P27470/D90216/A38175	4
	<i>S. downei (sobrinus)</i>	DEI	L34406	5
40	<i>S. mutants</i> (Ingbritt)	GBP	M30945/A37184	6
	<i>S. mutants</i> (GS-5)	GTF-B	A33128	7
	<i>S. mutants</i> (GS-5)	GTF-B	P08987/M17361/B33135	8
	<i>S. mutants</i>	GTF-B ^{3'} -ORF	P05427/C33135	8
45	<i>S. mutants</i> (GS-5)	GTF-C	P13470/M17361/M22054	9
	<i>S. mutants</i> (GS-5)	GTF-C	not available	10
	<i>S. mutants</i> (GS-5)	GTF-D	M29296/A45866	11
50	<i>S. salivarius</i>	GTF-J	A44811/S22726/S28809	12
	<i>S. salivarius</i>	GTF-K	Z11873/M64111	
	<i>S. salivarius</i> (ATCC25975)	GTF-L	S22737/S22727/Z11872	13
	<i>S. salivarius</i> (ATCC25975)	GTF-L	L35495	14
	<i>S. salivarius</i> (ATCC25975)	GTF-M	L35928	14
55	<i>S. pneumoniae</i> R6	LytA	P06653/A25634/M13812	15
	<i>S. pneumoniae</i>	PspA	A41971/M74122	16

5	Phage HB-3	HBL	P32762/M34652	17
	Phage Cp-1	CPL-1	P15057/J03586/A31086	18
	Phage Cp-9	CPL-9	P19386/M34780/JQ0438	19
	Phage EJ-1	EJL	A42936	20
10	<i>C. difficile</i> (VPI 10463)	ToxA	P16154/A37052/M30307	21
			X51797/S08638	
	<i>C. difficile</i> (BARTS W1)	ToxA	A60991/X17194	22
	<i>C. difficile</i> (VPI 10463)	ToxB	P18177/X53138/X60984	23,24
15			S10317	
	<i>C. difficile</i> (1470)	ToxB	S44271/Z23277	25,26
	<i>C. novyi</i>	a-toxin	S44272/Z23280	27
	<i>C. novyi</i>	a-toxin	Z48636	28
20	<i>C. acetobutylicum</i> (NCIB8052)	CspA	S49255/Z37723	29
	<i>C. acetobutylicum</i> (NCIB8052)	CspB	Z50008	30
	<i>C. acetobutylicum</i> (NCIB8052)	CspC	Z50033	30
	<i>C. acetobutylicum</i> (NCIB8052)	CspD	Z50009	30

2References:

- 1) Sato *et al.* (1993) *DNA sequence* 4, 19-27
 - 2) Ferreti *et al.* (1987) *J. Bacteriol.* 169, 4271-4278
 - 3) Gilmore *et al.* (1990) *J. Infect. Immun.* 58, 2452-2458
 - 4) Abo *et al.* (1991) *J. Bacteriol.* 173, 989-996
 - 5) Sun *et al.* (1994) *J. Bacteriol.* 176, 7213-7222
 - 6) Banas *et al.* (1990) *J. Infect. Immun.* 58, 667-673
 - 7) Shiroza *et al.* (1990) *Protein Sequence Database*
 - 8) Shiroza *et al.* (1987) *J. Bacteriol.* 169, 4263-4270
 - 9) Ueda *et al.* (1988) *Gene* 69, 101-109
 - 10) Russel (1990) *Arch. Oral. Biol.* 35, 53-58
 - 11) Honda *et al.* (1990) *J. Gen. Microbiol.* 136, 2099-2105
 - 12) Giffard *et al.* (1991) *J. Gen. Microbiol.* 137, 2577-2593
 - 13) Jacques (1992) *EMBL Data Library*
 - 14) Simpson *et al.* (1995) *J. Infect. Immun.* 63, 609-621
 - 15) Gargia *et al.* (1986) *Gene* 43, 265-272
 - 16) Yother *et al.* (1992) *J. Bacteriol.* 174, 601-609
 - 17) Romero *et al.* (1990) *J. Bacteriol.* 172, 5064-5070
 - 18) Garcia *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85, 914-918
 - 19) Garcia *et al.* (1990) *Gene* 86, 81-88
 - 20) Diaz *et al.* (1992) *J. Bacteriol.* 174, 5516-5525
 - 21) Dove *et al.* (1990) *J. Infect. Immun.* 58, 480-488
 - 22) Wren *et al.* (1990) *FEMS Microbiol. Lett.* 70, 1-6
 - 23) Barroso *et al.* (1990) *Nucleic Acids Res.* 18, 4004-4004
 - 24) von Eichel-Streiber *et al.* (1992) *Mol. Gen. Genet.* 233, 260-268
 - 25) Sartinger *et al.* (1993) *EMBL Data Library*
 - 26) von Eichel-Streiber *et al.* (1995) *Mol. Microbiol.* In Press
 - 27) Hofmann *et al.* (1993) *EMBL Data Library*
 - 28) Hofmann *et al.* (1995) *Mol. Gen. Genet.* In Press
 - 29) Sanchez *et al.* (1994) *EMBL Data Library*
 - 30) Sanchez *et al.* (1995) *EMBL Data Library*
- New cellulose binding peptides with interesting binding characteristics and specificities can be identified and screened for and the

genes encoding same isolated using well known molecular biology approaches combined with a variety of other procedures including, for example, spectroscopic (titration) methods such as: NMR spectroscopy (Zhu *et al.* Biochemistry (1995) 34:13196-13202, Gehring *et al.* Biochemistry (1991) 30:5524-5531), UV difference spectroscopy (Belshaw *et al.* Eur. J. Biochem. (1993) 211:717-724), fluorescence (titration) spectroscopy (Miller *et al.* J. Biol. Chem. (1983) 258:13665-13672), UV or fluorescence stopped flow analysis (De Boeck *et al.* Eur. J. Biochem. (1985) 149:141-415), affinity methods such as affinity electrophoresis (Mimura *et al.* J. chromatography (1992) 597:345-350) or affinity chromatography on immobilized mono or oligosaccharides, precipitation or agglutination analysis including turbidimetric or nephelometric analysis (Knibbs *et al.* J. Biol. Chem. (1993) 14940-14947), competitive inhibition assays (with or without quantitative IC50 determination) and various physical or physico-chemical methods including differential scanning or isothermal titration calorimetry (Sigurskjold *et al.* J. Biol. Chem. (1992) 267:8371-8376; Sigurskjold *et al.* Eur. J. Biol. (1994) 225:133-141) or comparative protein stability assays (melts) in the absence or presence of oligo saccharides using thermal CD or fluorescence spectroscopy.

The K_a for binding of the cellulose binding domains or proteins to cellulose is at least in the range of weak antibody-antigen extractions, i.e., $\geq 10^3$, preferably 10^4 , most preferably 10^6 M⁻¹. If the binding of the cellulose binding domain or protein to cellulose is exothermic or endothermic, then binding will increase or decrease, respectively, at lower temperatures, providing a means for temperature modulation of the binding step, see numeral 14 in Figure 1.

Table 3
Overview of enzymes with chitin-binding domains

Source (strain)	Enzyme	Accession No.	Ref. ³
Bacterial enzymes			
<u>Type I</u>			
<i>Aeromonas</i> sp. (No10S-24)	Chi	D31818	1
<i>Bacillus circulans</i> (WL-12)	ChiA1	P20533/M57601/A38368	2
<i>Bacillus circulans</i> (WL-12)	ChiD	P27050/D10594	3
<i>Janthinobacterium lividum</i>	Chi69	U07025	4

WO 00/77174

PCT/IL00/00330

	<i>Streptomyces griseus</i>	Protease C	21 A53669	5
	<u>Type II</u>			
5	<i>Aeromonas cavia</i> (K1)	Chi	U09139	6
	<i>Alteromonas</i> sp (0-7)	Chi85	A40633/P32823/D13762	7
	<i>Autographa californica</i> (C6)	NPH-128 ^a	P41684/L22858	8
	<i>Serratia marcescens</i>	ChiA	A25090/X03657/L01455/P07254	9
10	<u>Type III</u>			
	<i>Rhizopus oligosporus</i> (IFO8631)	Chi1	P29026/A47022/D10157/S27418	10
	<i>Rhizopus oligosporus</i> (IFO8631)	Chi2	P29027/B47022/D10158/S27419	10
15	<i>Saccharomyces cerevisiae</i>	Chi	S50371/U17243	11
	<i>Saccharomyces cerevisiae</i> Chi1 (DBY939)	P29028/M74069	12	
	<i>Saccharomyces cerevisiae</i> Chi2 (DBY918)	P29029/M7407/B41035	12	
20	Plant enzymes			
	<u>Hevein superfamily</u>			
25	<i>Allium sativum</i>	Chi	M94105	13
	<i>Amaranthus caudatus</i>	AMP-1 ^b	P27275/A40240	14, 15
	<i>Amaranthus caudatus</i>	AMP-2 ^b	S37381/A40240	14, 15
	<i>Arabidopsis thaliana</i> (cv. colombia)	ChiB	P19171/M38240/B45511	16
30	<i>Arabidopsis thaliana</i>	PHP ^c	U01880	17
	<i>Brassica napus</i>	Chi	U21848	18
	<i>Brassica napus</i>	Chi2	Q09023/M95835	19
	<i>Hevea brasiliensis</i>	Hev1 ^d	P02877/M36986/A03770/A38288	20, 21
	<i>Hordeum vulgare</i>	Chi33	L34211	22
35	<i>Lycopersicon esculentum</i>	Chi9	Q05538/Z15140/S37344	23
	<i>Nicotiana tabacum</i>	CBP20 ^e	S72424	24
	<i>Nicotiana tabacum</i>	Chi	A21091	25
	<i>Nicotiana tabacum</i> (cv. Havana)	Chi	A29074/M15173/S20981/S19855	26
	<i>Nicotiana tabacum</i> (FB7-1)	Chi	JQ0993/S0828	27
40	<i>Nicotiana tabacum</i> (cv. Samsun)	Chi	A16119	28
	<i>Nicotiana tabacum</i> (cv. Havana)	Chi	P08252/X16939/S08627	27
	<i>Nicotiana tabacum</i> (cv. BY4)	Chi	P24091/X51599/X64519/S13322	26, 27, 29
	<i>Nicotiana tabacum</i> (cv. Havana)	Chi	P29059/X64518/S20982	26
	<i>Oryza sativum</i> (IR36)	ChiA	L37289	30
45	<i>Oryza sativum</i>	ChiB	JC2253/S42829/Z29962	31
	<i>Oryza sativum</i>	Chi	S39979/S40414/X56787	32
	<i>Oryza sativum</i> (cv. Japonicum)	Chi	X56063	33
	<i>Oryza sativum</i> (cv. Japonicum)	Chi1	P24626/X54367/S14948	34
	<i>Oryza sativum</i>	Chi2	P25765/S15997	35
50	<i>Oryza sativum</i> (cv. Japonicum)	Chi3	D16223	
	<i>Oryza sativum</i>	ChiA	JC2252/S42828	30
	<i>Oryza sativum</i>	Chi1	D16221	32
	<i>Oryza sativum</i> (IR58)	Chi	U02286	36
	<i>Oryza sativum</i>	Chi	X87109	37
55	<i>Pisum sativum</i> (cv. Birte)	Chi	P36907/X63899	38
	<i>Pisum sativum</i> (cv. Alcan)	Chi2	L37876	39
	<i>Populus trichocarpa</i>	Chi	S18750/S18751/X59995/P29032	40
	<i>Populus trichocarpa</i> (H11-11)	Chi	U01660	41

WO 00/77174

PCT/IL00/00330

		22		
	<i>Phaseolus vulgaris</i> (cv. Saxa)	Chi	A24215/S43926/Jq0965/P36361	42
	<i>Phaseolus vulgaris</i> (cv. Saxa)	Chi	P06215/M13968/M19052/A25898	43,44,45
	<i>Sambucus nigra</i>	PR-3 ^f	Z46948	46
	<i>Secale cereale</i>	Chi	JC2071	47
5	<i>Solanum tuberosum</i>	ChiB1	U02605	48
	<i>Solanum tuberosum</i>	ChiB2	U02606	48
	<i>Solanum tuberosum</i>	ChiB3	U02607/S43317	48
	<i>Solanum tuberosum</i>	ChiB4	U02608	48
	<i>Solanum tuberosum</i>	WIN-1 ^g	P09761/X13497/S04926	49
10	(cv. Maris Piper)			
	<i>Solanum tuberosum</i>	WIN-2 ^g	P09762/X13497/S04927	49
	(cv. Maris Piper)			
	<i>Triticum aestivum</i>	Chi	S38670/X76041	50
	<i>Triticum aestivum</i>	WGA-1 ^h	P10968/M25536/S09623/S07289	51,52
15	<i>Triticum aestivum</i>	WGA-2 ^h	P02876/M25537/S09624	51,53
	<i>Triticum aestivum</i>	WGA-3	P10969/J02961/S10045/A28401	54
	<i>Ulmus americana</i> (NPS3-487)	Chi	L22032	55
	<i>Urtica dioica</i>	AGL ⁱ	M87302	56
	<i>Vigna unguiculata</i>	Chi1	X88800	57
20	(cv. Red caloonia)			

^aNHP : nuclear polyhedrosis virus endochitinase like sequence; Chi : chitinase, ^banti-microbial peptide, ^cpre-hevein like protein, ^dhevein, ^echitin-binding protein, ^fpathogenesis related protein, ^gwound-induced protein, ^hwheat germ agglutinin, ⁱagglutinin (lectin).

References:

- 1) Udea et al. (1994) *J. Ferment. Bioeng.* 78, 205-211
- 2) Watanabe et al. (1990) *J. Biol. Chem.* 265, 15659-16565
- 3) Watanabe et al. (1992) *J. Bacteriol.* 174, 408-414
- 4) Gleave et al. (1994) *EMBL Data Library*
- 5) Sidhu et al. (1994) *J. Biol. Chem.* 269, 20167-20171
- 6) Jones et al. (1986) *EMBO J.* 5, 467-473
- 7) Sitrit et al. (1994) *EMBL Data Library*
- 8) Genbank entry only
- 9) Tsujibo et al. (1993) *J. Bacteriol.* 175, 176-181
- 10) Yanai et al. (1992) *J. Bacteriol.* 174, 7398-7406
- 11) Pauley (1994) *EMBL Data Library*
- 12) Kuranda et al. (1991) *J. Biol. Chem.* 266, 19758-19767
- 13) van Damme et al. (1992) *EMBL Data Library*
- 14) Broekaert et al. (1992) *Biochemistry* 31, 4308-4314
- 15) de Bolle et al. (1993) *Plant Mol. Physiol.* 22, 1187-1190
- 16) Samac et al. (1990) *Plant Physiol.* 93, 907-914
- 17) Potter et al. (1993) *Mol. Plant Microbe Interact.* 6, 680-685
- 18) Buchanan-Wollaston (1995) *EMBL Data Library*
- 19) Hamel et al. (1993) *Plant Physiol.* 101, 1403-1403
- 20) Broekaert et al. (1990) *Proc. Natl. Acad. Sci. USA* 87, 7633-7637
- 21) Lee et al. (1991) *J. Biol. Chem.* 266, 15944-15948
- 22) Leah et al. (1994) *Plant Physiol.* 6, 579-589
- 23) Danhash et al. (1993) *Plant Mol. Biol.* 22 1017-1029
- 24) Ponstein et al. (1994) *Plant Physiol.* 104, 109-118
- 25) Meins et al. (1991) Patent EP0418695-A1
- 26) van Buuren et al. (1992) *Mol. Gen. Genet.* 232, 460-469
- 27) Shinshi et al. (1990) *Plant Mol. Biol.* 14, 357-368
- 28) Cornellisen et al. (1991) Patent EP0440304-A2
- 29) Fukuda et al. (1991) *Plant Mol. Biol.* 16, 1-10
- 30) Yun et al. (1994) *EMBL Data Library*
- 31) Kim et al. (1994) *Biosci. Biotechnol. Biochem.* 58, 1164-1166

- 32) Nishizawa *et al.* (1993) *Mol. Gen. Genet.* **241**, 1-10
 33) Nishizawa *et al.* (1991) *Plant Sci* **76**, 211-218
 34) Huang *et al.* (1991) *Plant Mol. Biol.* **16**, 479-480
 35) Zhu *et al.* (1991) *Mol. Gen. Genet.* **226**, 289-296
 5 36) Muthukrishnan *et al.* (1993) *EMBL Data Library*
 37) Xu (1995) *EMBL Data Library*
 38) Vad *et al.* (1993) *Plant Sci* **92**, 69-79
 39) Chang *et al.* (1994) *EMBL Data Library*
 40) Davis *et al.* (1991) *Plant Mol. Biol.* **17**, 631-639
 10 41) Clarke *et al.* (1994) *Plant Mol. Biol.* **25**, 799-815
 42) Broglie *et al.* (1989) *Plant Cell* **1**, 599-607
 43) Broglie *et al.* (1986) *Proc. Natl. acad. Sci. USA* **83**, 6820-6824
 44) Lucas *et al.* (1985) *FEBS Lett.* **193**, 208-210
 45) Hedrick *et al.* (1988) *Plant Physiol.* **86**, 182-186
 15 46) Roberts *et al.* (1994) *EMBL Data Library*
 47) Vamagami *et al.* (1994) *Biosci. Biotechnol. Biochem.* **58**, 322-329
 48) Beerhues *et al.* (1994) *Plant Mol. Biol.* **24**, 353-367
 49) Stanford *et al.* (1989) *Mol. Gen. Genet.* **215**, 200-208
 50) Liao *et al.* (1993) *EMBL Data Library*
 20 51) Smith *et al.* (1989) *Plant Mol. Biol.* **13**, 601-603
 52) Wright *et al.* (1989) *J. Mol. Evol.* **28**, 327-336
 53) Wright *et al.* (1984) *Biochemistry* **23**, 280-287
 54) Raikhel *et al.* (1987) *Proc. Natl. acad. Sci. USA* **84**, 6745-6749
 55) Hajela *et al.* (1993) *EMBL Data Library*
 25 56) Lerner *et al.* (1992) *J. Biol. Chem.* **267**, 11085-11091
 57) Vo *et al.* (1995) *EMBL Data Library*

Table 4
Sources of polysaccharide binding domains

30	Binding Domain	Proteins Where Binding Domain is Found
35	Cellulose Binding Domains ¹	β -glucanases (avicelases, CMCases, cellodextrinases) exoglucanases or cellobiohydrolases cellulose binding proteins xylanases mixed xylanases/glucanases esterases chitinases β -1,3-glucanases β -1,3-(β -1,4)-glucanases (β -)mannanases β -glucosidases/galactosidases cellulose synthases (unconfirmed)
50	Starch/Maltodextrin Binding Domains	α -amylases ^{2,3} β -amylases ^{4,5} pullulanases glucoamylases ^{6,7} cyclodextrin glucotransferases ⁸⁻¹⁰ (cyclomaltodextrin glucanotransferases) maltodextrin binding proteins ¹¹
55		

5	Dextran Binding Domains	(Streptococcal) glycosyl transferases ¹² dextran sucrases (unconfirmed) Clostridial toxins ^{13,14} glucoamylases ⁶ dextran binding proteins
10	β -Glucan Binding Domains	β -1,3-glucanases ^{15,16} β -1,3-(β -1,4)-glucanases (unconfirmed) β -1,3-glucan binding protein
15	Chitin Binding Domains	chitinases chitobiases chitin binding proteins (see also cellulose binding domains) Hevein

¹Gilkes *et al.*, *Adv. Microbiol Reviews*, (1991) 303-315.

²S?gaard *et al.*, *J. Biol. Chem.* (1993) 268:22480.

³Weselake *et al.*, *Cereal Chem.* (1983) 60:98.

⁴Svensson *et al.*, *J.* (1989) 264:309.

⁵Jespersen *et al.*, *J.* (1991) 280:51.

⁶Belshaw *et al.*, *Eur. J. Biochem.* (1993) 211:717.

⁷Sigurskjold *et al.*, *Eur. J. Biochem.* (1994) 225:133.

⁸Villette *et al.*, *Biotechnol. Appl. Biochem.* (1992) 16:57.

⁹Fukada *et al.*, *Biosci. Biotechnol. Biochem.* (1992) 56:556.

¹⁰Lawson *et al.*, *J. Mol. Biol.* (1994) 236:590.

¹⁴von Eichel-Streiber *et al.*, *Mol. Gen. Genet.* (1992) 233:260.

¹⁵Klebl *et al.*, *J. Bacteriol.* (1989) 171:6259.

¹⁶Watanabe *et al.*, *J. Bacteriol.* (1992) 174:186.

¹⁷Duvic *et al.*, *J. Biol. Chem.* (1990) :9327.

Thus, and as already stated, the phrase "polysaccharide binding peptide" includes an amino acid sequence which comprises at least a functional portion of a polysaccharide binding region (domain) of a polysaccharidase or a polysaccharide binding protein. The phrase further relates to a polypeptide screened for its cellulose binding activity out of a library, such as a peptide library or a DNA library (e.g., a cDNA library or a display library). By "functional portion" is intended an amino acid sequence which binds to cellulose.

The techniques used in isolating polysaccharidase genes, such as cellulase genes, and genes for cellulose binding proteins are known in the art, including synthesis, isolation from genomic DNA, preparation from cDNA, or combinations thereof. (See, U.S. Pat. Nos. 5,137,819; 5,202,247; 5,340,731; 5,496,934; and 5,837,814). The sequences for several binding domains, which bind to soluble oligosaccharides are known (See, Figure 1 of PCT/CA97/00033, WO 97/26358). The DNAs coding for a variety of polysaccharidases and polysaccharide binding proteins are also known. Various techniques for manipulation of genes are well known, and

include restriction, digestion, resection, ligation, *in vitro* mutagenesis, primer repair, employing linkers and adapters, and the like (see Sambrook *et al.*, Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989, which is incorporated
5 herein by reference).

The amino acid sequence of a polysaccharidase can be used to design a probe to screen a cDNA or a genomic library prepared from mRNA or DNA from cells of interest as donor cells for a polysaccharidase gene or a polysaccharide binding protein gene. By using the polysaccharidase cDNA
10 or binding protein cDNA or a fragment thereof as a hybridization probe, structurally related genes found in other species can be easily cloned and provide a cellulose binding peptide which is expressible in plants according to the present invention. Particularly contemplated is the isolation of genes from organisms that express polysaccharidase activity using oligonucleotide
15 probes based on the nucleotide sequences of genes obtainable from an organism wherein the catalytic and binding domains of the polysaccharidase are discrete, although other polysaccharide binding proteins also can be used (see, for example, Shoseyov, *et al.*, Proc. Nat'l. Acad. Sci. (USA) (1992) 89:3483-3487).

Probes developed using consensus sequences for the binding domain of a polysaccharidase or polysaccharide-binding protein are of particular interest. The β -1,4-glycanases from *C. fimi* characterized to date are endoglucanases A, B, C and D (CenA, CenB, CenC and CenD, respectively), exocellobiohydrolases A and B (CbhA and CbhB,
25 respectively), and xylanases A and D (Cex and XylD, respectively) (see Wong *et al.* (1986) Gene, 44:315; Meinke *et al.* (1991) J. Bacteriol., 173:308; Coutinho *et al.*, (1991) Mol. Microbiol. 5:1221; Meinke *et al.*, (1993) Bacteriol., 175:1910; Meinke *et al.*, (1994) Mol. Microbiol., 12:413; Shen *et al.*, Biochem. J., in press; O'Neill *et al.*, (1986) Gene, 44:325; and
30 Millward-Sadler *et al.*, (1994) Mol. Microbiol., 11:375). All are modular proteins of varying degrees of complexity, but with two features in common: a catalytic domain (CD) and a cellulose-binding domain (CBD) which can function independently (see Millward-Sadler *et al.*, (1994) Mol. Microbiol., 11:375; Gilkes *et al.*, (1988) J. Biol. Chem., 263:10401; Meinke *et al.*, (1991) J. Bacteriol., 173:7126; and Coutinho *et al.*, (1992) Mol. Microbiol., 6:1242). In four of the enzymes, CenB, CenD, CbhA and CbhB, fibronectin type III (Fn3) repeats separate the N-terminal CD from the C-terminal CBD. The CDs of the enzymes come from six of the

families of glycoside hydrolases (see Henrissat (1991) Biochem. J., 280:309; and Henrissat *et al.*, (1993) Biochem. J., 293:781); all of the enzymes have an N- or C-terminal CBD or CBDs (see Tomme *et al.*, Adv. Microb. Physiol., in press); CenC has tandem CBDs from family IV at its

5 N-terminus; CenB and XylD each have a second, internal CBD from families III and II, respectively. Cex and XylD are clearly xylanases; however, Cex, but not XylD, has low activity on cellulose. Nonetheless, like several other bacterial xylanases (see Gilbert *et al.*, (1993) J. Gen. Microbiol., 139:187), they have CBDs. *C. fimi* probably produces other β -

10 1,4-glycanases. Similar systems are produced by related bacteria (see Wilson (1992) Crit. Rev. Biotechnol., 12:45; and Hazlewood *et al.*, (1992) J. Appl. Bacteriol., 72:244). Unrelated bacteria also produce glycanases; *Clostridium thermocellum*, for example, produces twenty or more β -1,4-glycanases (see Beguin *et al.*, (1992) FEMS Microbiol. Lett., 100:523).

15 The CBD derived from *C. fimi* endoglucanase C N1, is the only protein known to bind soluble cellosaccharides and one of a small set of proteins that are known to bind any soluble polysaccharides.

Examples of suitable binding domains are shown in Figure 1 of PCT/CA97/00033 (WO 97/26358), which presents an alignment of binding

20 domains from various enzymes that bind to polysaccharides and identifies amino acid residues that are conserved among most or all of the enzymes. This information can be used to derive a suitable oligonucleotide probe using methods known to those of skill in the art. The probes can be considerably shorter than the entire sequence but should at least be 10,

25 preferably at least 14, nucleotides in length. Longer oligonucleotides are useful, up to the full length of the gene, preferably no more than 500, more preferably no more than 250, nucleotides in length. RNA or DNA probes can be used. In use, the probes are typically labeled in a detectable manner, for example, with ^{32}P , ^3H , biotin, avidin or other detectable reagents, and

30 are incubated with single-stranded DNA or RNA from the organism in which a gene is being sought. Hybridization is detected by means of the label after the unhybridized probe has been separated from the hybridized probe. The hybridized probe is typically immobilized on a solid matrix such as nitrocellulose paper. Hybridization techniques suitable for use with

35 oligonucleotides are well known to those skilled in the art. Although probes are normally used with a detectable label that allows easy identification, unlabeled oligonucleotides are also useful, both as precursors of labeled probes and for use in methods that provide for direct detection of double-

Generally, the binding domains identified by probing nucleic acids from an organism of interest will show at least about 40 % identity (including as appropriate allowances for conservative substitutions, gaps for better alignment and the like) to the binding region or regions from which the probe was derived and will bind to a soluble β -1,4 glucan with a K_a of $\geq 10^3 \text{ M}^{-1}$. More preferably, the binding domains will be at least about 60 % identical, and most preferably at least about 70 % identical to the binding region used to derive the probe. The percentage of identity will be greater among those amino acids that are conserved among polysaccharidase binding domains. Analyses of amino acid sequence comparisons can be performed using programs in PC/Gene (IntelliGenetics, Inc.). PCLUSTAL can be used for multiple sequence alignment and generation of phylogenetic trees.

In order to isolate the polysaccharide binding protein or a polysaccharide binding domain from an enzyme or a cluster of enzymes that binds to a polysaccharide, several genetic approaches can be used. One method uses restriction enzymes to remove a portion of the gene that codes for portions of the protein other than the binding portion thereof. The remaining gene fragments are fused with expression control sequences to obtain a mutated gene that encodes a truncated protein. Another method involves the use of exonucleases such as *Bal31* to systematically delete nucleotides either externally from the 5' and the 3' ends of the DNA or internally from a restricted gap within the gene. These gene deletion methods result in a mutated gene encoding a shortened protein molecule which can then be evaluated for substrate or polysaccharide binding ability.

Any cellulose binding protein or cellulose binding domain may be used in the present invention. The term "cellulose binding protein" ("CBP") refers to any protein or polypeptide which specifically binds to cellulose. The cellulose binding protein may or may not have cellulose or cellulolytic activity. The term "cellulose binding domain" ("CBD") refers to any protein or polypeptide which is a region or portion of a larger protein, said region or portion binds specifically to cellulose. The cellulose binding domain (CBD) may be a part or portion of a cellulase, xylanase or other polysaccharidase, e.g., a chitinase, etc., a sugar binding protein such as maltose binding protein, or scaffoldin such as CbpA of *Clostridium cellovorans*, etc. Many cellulases and hemicellulases (e.g. xylanases and mannanases) have the

ability to associate with cellulose. These enzymes typically have a catalytic domain containing the active site for substrate hydrolysis and a carbohydrate-binding domain or cellulose-binding domain for binding cellulose. The CBD may also be from a non-catalytic polysaccharide binding protein. To date, more than one hundred cellulose-binding domains (CBDs) have been classified into at least thirteen families designated I-XIII (Tomme *et al.* (1995) "Cellulose Binding Domains: Classification and Properties", in ACS Symposium Series 618 Enzymatic Degradation and Insoluble Carbohydrates, pp. 142-161, Saddler and Penner eds., American Chemical Society, Washington, D.C. (Tomme I) ; Tomme *et al.* Adv. Microb. Physiol. (1995) 37:1 (Tomme II); and Smant *et al.*, Proc. Natl. Acad. Sci U.S.A. (1998) 95:4906-4911, all of which are incorporated herein by reference). Any of the CBDs described in Tomme I or II or any variants thereof, any other presently known CBDs or any new CBDs which may be identified can be used in the present invention. As an illustrative, but in no way limiting example, the CBP or CBD can be from a bacterial, fungal, slime mold, or nematode protein or polypeptide. For a more particular illustrative example, the CBD is obtainable from *Clostridium cellulovorans*, *Clostridium cellulovorans*, or *Cellulomonas fimi* (e.g., CenA, CenB, CenD, Cex). In addition, the CBD may be selected from a phage display peptide or peptidomimetic library, random or otherwise, using e.g., cellulose as a screening agent. (See Smith Science (1985) 228:1315-1317 and Lam, Nature (1991) 354:82-84). Furthermore, the CBD may be derived by mutation of a portion of a protein or polypeptide which binds to a polysaccharide other than cellulose (or hemicellulose) but also binds cellulose, such as a chitinase, which specifically binds chitin, or a sugar binding protein such as maltose binding protein, rendering said portion capable of binding to cellulose. In any event, the CBD binds cellulose or hemicellulose. Shoseyov and Doi (Proc. Natl. Acad. Sci. USA (1990) 87:2192-2195) isolated a unique cellulose-binding protein (CbpA) from the cellulose "complex" of the cellulolytic bacterium *Clostridium cellulovorans*. This major subunit of the cellulose complex was found to bind to cellulose, but had no hydrolytic activity, and was essential for the degradation of crystalline cellulose. The CbpA gene has been cloned and sequenced (Shoseyov *et al.* Proc. Natl. Acad. Sci. USA (1992) 89:3483-3487). Using PCR primers flanking the cellulose-binding domain of CbpA, the latter was successfully cloned into an overexpression vector that enabled overproduction of the approximately 17 kDa CBD in *Escherichia coli*. The

recombinant CBD exhibits very strong affinity to cellulose and chitin (U.S. Pat. No. 5,496,934; Goldstein *et al.*, J. Bacteriol. (1993) 175:5762; PCT International Publication WO 94/24158, all are incorporated by reference as if fully set forth herein).

5 In recent years, several CBDs have been isolated from different sources. Most of these have been isolated from proteins that have separate catalytic, i.e., cellulose and cellulose binding domains, and only two have been isolated from proteins that have no apparent hydrolytic activity but possess cellulose-binding activity (Goldstein *et al.* J. Bacteriol. (1993)
10 175:5762-5768; Morag *et al.* Appl. (1995) Environ. Microbiol. 61:1980-1986).

Recombinant proteins:

Any protein for which a gene is known or can be isolated can be used as the recombinant protein and be fused to the cellulose binding peptide
15 according to the present invention. Advantageously, the recombinant protein is of a commercial value. A non-exhaustive list of recombinant proteins which can be manufactured utilizing the process of the present invention and their uses follows.

20 Thus, for example, glucoamylases and glucose isomerases are used in the food processing industry to convert starch to high fructose corn syrup.

Another useful class of enzymes are proteinases, which are used for the hydrolysis of high molecular weight proteins and which are further used in combination with detergents in cleaning applications, in leather manufacturing processes, in the food industry, and in the manufacture of
25 alcoholic beverages.

Enzymes known as pectinesterases, and several related enzymes, are used for pectin hydrolysis in the food industry.

A class of enzymes known as lipases are used for the cleavage of ester linkages in triglycerides, and are used both in the food industry and for
30 effluent treatment.

The enzyme beta-galactosidase is used industrially for the hydrolysis of whey lactose.

An enzyme known as thermolysin is used in the production of the artificial sweetener aspartame.

35 An enzyme known as sulphydryl oxidase is used in the reduction of the cooked flavor of milk.

Enzymes known as catalases are used to remove hydrogen peroxides from milk, cheese, and egg processing, and are further used in the sterilization and oxidation of plastics and rubbers.

5 Heparinases are useful for the production of heparin and heparan sulfate oligosaccharides.

Other proteins, in addition to enzymes, are those which have affinities to other compounds. For example, bacteria, fungi, plants and animals all contain a large number of proteins that exhibit specific interactions with agents such as metal ions and toxic compounds, and have
10 high affinities for such agents.

A class of proteins known as metalloproteins contain prosthetic groups that bind specifically to metal ions. An example of such a prosthetic group is the porphyrin group in hemoglobin. Some other examples of metal ion binding proteins include parvalbumin, which binds to calcium, and
15 metallothionin, an animal protein that binds large amounts of metal ions, especially zinc. Such metal absorptive proteins could also be used for purification in industrial processes.

It is also envisioned that streams of flowing material could be degraded by microbial enzymes. It is known that certain pollutants, whether
20 natural or synthetic, and certain pesticides and other durable organic compounds in the environment can be degraded (inactivated) or converted into useful compounds by microbial enzymes.

It is known, for example, that some microorganisms, for example *Pseudomonas putida*, possessed dehalogenases that are capable of
25 degrading certain pesticides and herbicides, and rendering them less toxic. Similarly, hydrolysis of organophosphate insecticides have been observed by microbial enzymes.

It is also possible to produce antibodies within plant cells. The antibodies can include monoclonal antibodies or fragments thereof having
30 at least a portion of an antigen binding region, including immunoactive entities such as Fv, F(abl)2, Fab fragments (Harlow and Lane, 1988 Antibody, Cold Spring Harbor), single chain antibodies (U.S. Pat. No. 4,946,778), chimeric or humanized antibodies (Morrison *et al.* Proc. Natl. Acad. Sci. USA (1984) 81:6851; Neuberger *et al.* Nature (1984) 312:604-
35 608) and complementarily determining regions (CDR).

Another class of proteins are those that bind to antibodies, such as protein-A, protein-G, protein-L and their mutants.

It is also possible to produce protein antibiotics or peptides such as lysozyme or therapeutic proteins which might assist in healing processes, for example, certain wound healing peptides, growth factors and hormones. Proteins such as HSA can also be produced.

5 Another class of proteins include proteins such as agglutinin, zein, silk, elastine proteins as well as COMP, JUN, FOS and other proteins that may form stable protein-protein interactions such as coiled-coil interactions that may be useful for production of protein fibers.

Another example is the production of animal feed enzymes. Phytase
10 from *Aspergillus niger*, for example, increases the availability of phosphorus from feed for monogastric animals by releasing phosphate from the substrate phytic acid, therefore reducing the need for costly phosphorus supplements. A phytase cDNA was constitutively expressed in transgenic tobacco (*Nicotiana tabacum*) plants (Verwoerd *et al.*, Plant. Physiol.
15 (1995) 109:1199-205). Soybean plants transformed with a fungal phytase gene improve phosphorus availability whereas excretion was decreased for broilers. It appears that phytase can improve growth performance of broilers fed low phosphorous diets when provided either as a commercial supplement or in the form of transformed seeds (Denbow *et al.*, Poult. Sci.
20 (1989) 77:878-881).

Other recombinant proteins of interest, will for the most part be mammalian proteins, and will include blood proteins, such as serum albumin, Factor VII, Factor VIIIc, Factor VIIIvW, Factor IX, Factor X, tissue plasminogen factor, Protein C, von Willebrand factor, antithrombin
25 III, erythropoietin, colony stimulating factors, such as G-, M-, GM-, cytokines, such as interleukins 1-11, integrins, addressins, selectins, homing receptors, surface membrane proteins, such as surface membrane protein receptors, T cell receptor units, immunoglobulins (as further detailed above with respect to antibodies), soluble major histocompatibility complex
30 antigens, structural proteins, such as collagen, fibrin, elastin, tubulin, actin, and myosin, growth factor receptors, growth factors, growth hormone, cell cycle proteins, vaccines, fibrinogen, thrombin, cytokines and hyaluronidase. Additional examples include chymosin, polymerases, saccharidases, dehydrogenases, nucleases, oxido reductases such as fungal peroxidases and
35 lactases, xylanases, rennin, horse radish peroxidase, amylases and soil remediation enzymes.

The genes encoding all of the above listed proteins have been isolated and as such these proteins are readily available for recombinant

expression and production according to the teachings of the present invention. It will be appreciated that new genes encoding an ever growing spectrum of proteins are continuously discovered and isolated, rendering such genes available for molecular manipulation and recombinant
5 expression. There is thus no intention to limit the recombinant protein produced utilizing the method of the present invention to any specific protein or list of proteins.

Cellulose binding peptide-recombinant protein fusions:

The fusion of two proteins for which genes has been isolated is well
10 known and practiced in the art. Such fusion involves the joining together of heterologous nucleic acid sequences, in frame, such that translation thereof results in the generation of a fused protein product or a fusion proteins. Methods, such as the polymerase chain reaction (PCR), restriction, nuclease digestion, ligation, synthetic oligonucleotides synthesis and the like are
15 typically employed in various combinations in the process of generating fusion gene constructs. One ordinarily skilled in the art can readily form such constructs for any pair or more of individual proteins. Interestingly, in most cases where such fusion or chimera proteins are produced, and in all cases where one of the proteins was a cellulose binding peptide, both the
20 former and the latter retained their catalytic activity or function.

For example, Greenwood *et al.* (1989, FEBS Lett. 224:127-131) fused the cellulose binding region of *Cellulomonas fimi* endoglucanase to the enzyme alkaline phosphatase. The recombinant fusion protein retained both its phosphatase activity and the ability to bind to cellulose. For more
25 descriptions of cellulose binding fusion proteins, see U.S. Patent No. 5,137,819 issued to Kilburn *et al.*, and U.S. Patent No. 5,719,044 issued to Shoseyov *et al.* both incorporated by reference herein. See also U.S. Pat. No. 5,474,925. All of which are incorporated herein by reference.

The recombinant protein immobilized via its fused counterpart to the
30 cellulosic matter can be released from the plant derived cellulosic matter by cleavage thereof, e.g., by proteolysis, using either a nonspecific general protease such as proteinase K or trypsin, or a specific protease as further detailed hereinunder. For example, release can be effected by treatment with proteinase K at a concentration of about 50 µg/ml for about 20 minutes
35 at about 37 °C (Din *et al.* Bio/Technology (1991) 9:1096-1099).

Inclusion of a dedicated cleavage site:

According to a preferred embodiment of the present invention the fusion protein includes the recombinant protein and the cellulose binding

peptide separated therebetween via a unique amino acid sequence recognizable and digestible by a protease or under predetermined cis or trans conditions for digesting a controllable intervening protein sequence.

As used herein in the specification and in the claims section that follows, the phrase "unique amino acid sequence recognizable and digestible by a protease" includes a protease recognition sequence which is both recognizable and readily accessible to a protease. Thus, the unique sequence can be a solitary sequence (i.e., which does not appear in the recombinant protein and optionally also not in the cellulose binding peptide) or alternatively, the sole sequence of several similar sequences which is not sequestered from the protease due to the tertiary structure of the recombinant protein and optionally the cellulose binding peptide. In both these cases proteolysis will release the recombinant protein from the fusion protein cellulosic matter complex.

As used herein in the specification and in the claims section that follows, the phrase "controllable intervening protein sequence" includes unique amino acid sequences capable of excision from or cleavage of a peptide under predetermined conditions in cis or in trans, e.g., increase in temperature, exposure to light, unblocking of amino acid residues by dephosphorylation and treatment with chemical reagents or deglycosylation, examples include proteolysis effected under acidic conditions (HCl, e.g., to cleave between Asp and Pro) and proteolysis effected by a proteolyzing reagent, such as CNBr to cleave downstream of Met, all as known in the art and/or as further described in U.S. Pat. No. 5,834,247, which is incorporated herein by reference.

Thus, according to an aspect of the present invention there is provided a composition of matter comprising (a) a plant derived cellulosic matter; and (b) a fusion protein including a recombinant protein and a cellulose binding peptide separated therebetween via a unique amino acid sequence recognizable and digestible by a protease or under predetermined cis or trans conditions for digesting a controllable intervening protein sequence, wherein the fusion protein is complexed to the plant derived cellulosic matter by affinity binding via the cellulose binding peptide.

Nucleic acid molecules which can be used according to preferred embodiments of the present invention to express the fusion protein in plant cells would therefore include a heterologous nucleic acid sequence including (i) a first sequence encoding a cellulose binding peptide; (ii) a second sequence encoding a recombinant protein, wherein the first and

second sequences are joined together in frame in either orientation; and (iii) a third sequence encoding a unique amino acid sequence recognizable and digestible by a protease or under predetermined cis or trans conditions for digesting a controllable intervening protein sequence, the third sequence is
5 between and in frame with the first and second sequences.

Thus, specific cleavage can be used to release the recombinant protein from the fusion protein cellulosic matter complex. For example, one can include a protease recognition site or a chemical cleavage site between the recombinant protein and the cellulose binding peptide. Examples of
10 recognition sites include those for collagenase, thrombin, enterokinase, and Factor X_a which are cleaved specifically by the respective enzymes. Chemical cleavage sites sensitive, for example, to low pH or cyanogen bromide, can also be used.

Where cleavage is used, the recombinant protein can be cleaved
15 readily from the cellulosic matter by the use of a protease specific for a sequence present therebetween and the cellulose binding peptide.

It will be appreciated in this respect that four main classes of specific proteases are known, including (i) cysteine proteases, including cathepsin B and L; (ii) aspartyl protease cathepsin D; (iii) serine proteases including
20 plasmin, tissue-type plasminogen activator (tPA) and urokinase-type plasminogen activator (uPA), and (iv) matrix metalloproteinases (MMPs), including collagenases, gelatinases A and B (MMP2 and MMP9) and stromelysin (MMP3). Members of these protease families are commercially available and their recognition sequences known. As such, these proteases
25 can be used to implement the step of releasing the recombinant protein from the plant derived cellulosic matter while implementing the process according to the present invention.

Genetically modified plant material:

According to an aspect of the present invention there is provided a
30 nucleic acid molecule comprising (a) a promoter sequence for directing protein expression in plant cells; and (b) a heterologous nucleic acid sequence as further detailed herein, wherein, the heterologous nucleic acid sequence is down stream the promoter sequence, such that expression of the heterologous nucleic acid sequence is effectable by the promoter sequence.
35 Such a nucleic acid molecule needs to be effectively introduced into plant cells, so as to genetically modify the plant.

There are various methods of introducing foreign genes into both monocotyledonous and dicotyledonous plants (Potrykus, I., Annu. Rev.

Plant. Physiol., Plant. Mol. Biol. (1991) 42:205-225; Shimamoto *et al.*, Nature (1989) 338:274-276). The principle methods of causing stable integration of exogenous DNA into plant genomic DNA include two main approaches:

- 5 (i) *Agrobacterium*-mediated gene transfer: Klee *et al.* (1987)
Annu. Rev. Plant Physiol. 38:467-486; Klee and Rogers in Cell Culture
and Somatic Cell Genetics of Plants, Vol. 6, Molecular Biology of Plant
Nuclear Genes, eds. Schell, J., and Vařil, L. K., Academic Publishers, San
Diego, Calif. (1989) p. 2-25; Gatenby, in Plant Biotechnology, eds. Kung,
10 S. and Arntzen, C. J., Butterworth Publishers, Boston, Mass. (1989) p.
93-112.

- (ii) direct DNA uptake: Paszkowski *et al.*, in Cell Culture and Somatic Cell Genetics of Plants, Vol. 6, Molecular Biology of Plant Nuclear Genes eds. Schell, J., and Vasil, L. K., Academic Publishers, San Diego, Calif. (1989) p. 52-68; including methods for direct uptake of DNA into protoplasts, Toriyama, K. *et al.* (1988) Bio/Technology 6:1072-1074. DNA uptake induced by brief electric shock of plant cells: Zhang *et al.* Plant Cell Rep. (1988) 7:379-384. Fromm *et al.* Nature (1986) 319:791-793. DNA injection into plant cells or tissues by particle bombardment, Klein *et al.* Bio/Technology (1988) 6:559-563; McCabe *et al.* Bio/Technology (1988) 6:923-926; Sanford, Physiol. Plant. (1990) 79:206-209; by the use of micropipette systems: Neuhaus *et al.*, Theor. Appl. Genet. (1987) 75:30-36; Neuhaus and Spangenberg, Physiol. Plant. (1990) 79:213-217; or by the direct incubation of DNA with germinating pollen, DeWet *et al.* in Experimental Manipulation of Ovule Tissue, eds. Chapman, G. P. and Mantell, S. H. and Daniels, W. Longman, London, (1985) p. 197-209; and Ohta, Proc. Natl. Acad. Sci. USA (1986) 83:715-719.

The *Agrobacterium* system includes the use of plasmid vectors that contain defined DNA segments that integrate into the plant genomic DNA. Methods of inoculation of the plant tissue vary depending upon the plant species and the *Agrobacterium* delivery system. A widely used approach is the leaf disc procedure which can be performed with any tissue explant that provides a good source for initiation of whole plant differentiation. Horsch *et al.* in Plant Molecular Biology Manual A5, Kluwer Academic Publishers, Dordrecht (1988) p. 1-9. The *Agrobacterium* system is especially viable in the creation of transgenic dicotyledenous plants.

There are various methods of direct DNA transfer into plant cells. In electroporation, the protoplasts are briefly exposed to a strong electric field. In microinjection, the DNA is mechanically injected directly into the cells using very small micropipettes. In microparticle bombardment, the DNA is adsorbed on microprojectiles such as magnesium sulfate crystals or tungsten particles, and the microprojectiles are physically accelerated into cells or plant tissues.

Following transformation plant propagation is exercised. The most common method of plant propagation is by seed. Regeneration by seed propagation, however, has the deficiency that due to heterozygosity there is a lack of uniformity in the crop, since seeds are produced by plants according to the genetic variances governed by Mendelian rules. Basically, each seed is genetically different and each will grow with its own specific traits. Therefore, it is preferred that the transgenic plant be produced such that the regenerated plant has the identical traits and characteristics of the parent transgenic plant, e.g., a reproduction of the fusion protein. Therefore, it is preferred that the transgenic plant be regenerated by micropropagation which provides a rapid, consistent reproduction of the transgenic plants.

Micropropagation is a process of growing new generation plants from a single piece of tissue that has been excised from a selected parent plant or cultivar. This process permits the mass reproduction of plants having the preferred tissue expressing the fusion protein. The new generation plants which are produced are genetically identical to, and have all of the characteristics of, the original plant. Micropropagation allows mass production of quality plant material in a short period of time and offers a rapid multiplication of selected cultivars in the preservation of the characteristics of the original transgenic or transformed plant. The advantages of cloning plants are the speed of plant multiplication and the quality and uniformity of plants produced.

Micropropagation is a multi-stage procedure that requires alteration of culture medium or growth conditions between stages. Thus, the micropropagation process involves four basic stages: Stage one, initial tissue culturing; stage two, tissue culture multiplication; stage three, differentiation and plant formation; and stage four, greenhouse culturing and hardening. During stage one, initial tissue culturing, the tissue culture is established and certified contaminant-free. During stage two, the initial tissue culture is multiplied until a sufficient number of tissue samples are

produced to meet production goals. During stage three, the tissue samples grown in stage two are divided and grown into individual plantlets. At stage four, the transgenic plantlets are transferred to a greenhouse for hardening where the plants' tolerance to light is gradually increased so that
5 it can be grown in the natural environment.

The basic bacterial/plant vector construct will preferably provide a broad host range prokaryote replication origin; a prokaryote selectable marker; and, for *Agrobacterium* transformations, T DNA sequences for *Agrobacterium*-mediated transfer to plant chromosomes. Where the
10 heterologous sequence is not readily amenable to detection, the construct will preferably also have a selectable marker gene suitable for determining if a plant cell has been transformed. A general review of suitable markers for the members of the grass family is found in Wilmink and Dons, Plant Mol. Biol. Repr. (1993) 11:165-185.

15 Sequences suitable for permitting integration of the heterologous sequence into the plant genome are also recommended. These might include transposon sequences and the like for homologous recombination as well as Ti sequences which permit random insertion of a heterologous expression cassette into a plant genome.

20 Suitable prokaryote selectable markers include resistance toward antibiotics such as ampicillin or tetracycline. Other DNA sequences encoding additional functions may also be present in the vector, as is known in the art.

The constructs of the subject invention will include an expression
25 cassette for expression of the fusion protein of interest. Usually, there will be only one expression cassette, although two or more are feasible. The recombinant expression cassette will contain in addition to the heterologous sequence one or more of the following sequence elements, a promoter region, plant 5' untranslated sequences, initiation codon depending upon
30 whether or not the structural gene comes equipped with one, and a transcription and translation termination sequence. Unique restriction enzyme sites at the 5' and 3' ends of the cassette allow for easy insertion into a pre-existing vector.

Viral infected plant material:

35 Viruses are a unique class of infectious agents whose distinctive features are their simple organization and their mechanism of replication. In fact, a complete viral particle, or virion, may be regarded mainly as a block of genetic material (either DNA or RNA) capable of autonomous

replication, surrounded by a protein coat and sometimes by an additional membranous envelope such as in the case of alpha viruses. The coat protects the virus from the environment and serves as a vehicle for transmission from one host cell to another.

5 Viruses that have been shown to be useful for the transformation of plant hosts include CaV, TMV and BV. Transformation of plants using plant viruses is described in U.S. Pat. No. 4,855,237 (BGV), EP-A 67,553 (TMV), Japanese Published Application No. 63-14693 (TMV), EPA 194,809 (BV), EPA 278,667 (BV); and Gluzman, Y. *et al.*,
10 Communications in Molecular Biology: Viral Vectors, Cold Spring Harbor Laboratory, New York, pp. 172-189 (1988). Pseudovirus particles for use in expressing foreign DNA in many hosts, including plants, is described in WO 87/06261.

Construction of plant RNA viruses for the introduction and
15 expression of non-viral foreign genes in plants is demonstrated by the above references as well as by Dawson, W. O. *et al.*, Virology (1989) 172:285-292; Takamatsu *et al.* EMBO J. (1987) 6:307-311; French *et al.* Science (1986) 231:1294-1297; and Takamatsu *et al.* FEBS Letters (1990) 269:73-76.

20 When the virus is a DNA virus, the constructions can be made to the virus itself. Alternatively, the virus can first be cloned into a bacterial plasmid for ease of constructing the desired viral vector with the foreign DNA. The virus can then be excised from the plasmid. If the virus is a DNA virus, a bacterial origin of replication can be attached to the viral
25 DNA, which is then replicated by the bacteria. Transcription and translation of this DNA will produce the coat protein which will encapsidate the viral DNA. If the virus is an RNA virus, the virus is generally cloned as a cDNA and inserted into a plasmid. The plasmid is then used to make all of the constructions. The RNA virus is then produced by transcribing the viral
30 sequence of the plasmid and translation of the viral genes to produce the coat protein(s) which encapsidate the viral RNA.

Construction of plant RNA viruses for the introduction and expression of non-viral foreign genes in plants is demonstrated by the above references as well as in U.S. Pat. No. 5,316,931

35 In one embodiment, a plant viral nucleic acid is provided in which the native coat protein coding sequence has been deleted from a viral nucleic acid, a non-native plant viral coat protein coding sequence and a non-native promoter, preferably the subgenomic promoter of the non-native

coat protein coding sequence, capable of expression in the plant host, packaging of the recombinant plant viral nucleic acid, and ensuring a systemic infection of the host by the recombinant plant viral nucleic acid, has been inserted. Alternatively, the coat protein gene may be inactivated
5 by insertion of the non-native nucleic acid sequence within it, such that a fusion protein is produced. The recombinant plant viral nucleic acid may contain one or more additional non-native subgenomic promoters. Each non-native subgenomic promoter is capable of transcribing or expressing adjacent genes or nucleic acid sequences in the plant host and incapable of
10 recombination with each other and with native subgenomic promoters. Non-native (foreign) nucleic acid sequences may be inserted adjacent the native plant viral subgenomic promoter or the native and a non-native plant viral subgenomic promoters if more than one nucleic acid sequence is included. The non-native nucleic acid sequences are transcribed or
15 expressed in the host plant under control of the subgenomic promoter to produce the desired products.

In a second embodiment, a recombinant plant viral nucleic acid is provided as in the first embodiment except that the native coat protein coding sequence is placed adjacent one of the non-native coat protein
20 subgenomic promoters instead of a non-native coat protein coding sequence.

In a third embodiment, a recombinant plant viral nucleic acid is provided in which the native coat protein gene is adjacent its subgenomic promoter and one or more non-native subgenomic promoters have been
25 inserted into the viral nucleic acid. The inserted non-native subgenomic promoters are capable of transcribing or expressing adjacent genes in a plant host and are incapable of recombination with each other and with native subgenomic promoters. Non-native nucleic acid sequences may be inserted adjacent the non-native subgenomic plant viral promoters such that
30 said sequences are transcribed or expressed in the host plant under control of the subgenomic promoters to produce the desired product.

In a fourth embodiment, a recombinant plant viral nucleic acid is provided as in the third embodiment except that the native coat protein coding sequence is replaced by a non-native coat protein coding sequence.

35 The viral vectors are encapsidated by the coat proteins encoded by the recombinant plant viral nucleic acid to produce a recombinant plant virus. The recombinant plant viral nucleic acid or recombinant plant virus is used to infect appropriate host plants. The recombinant plant viral

nucleic acid is capable of replication in the host, systemic spread in the host, and transcription or expression of foreign gene(s) in the host to produce the desired fusion protein.

Fusion protein compartmentalization - signal peptides:

5 As already mentioned hereinabove, compartmentalization of the fusion protein is an important feature of the present invention because it allows undisturbed plant growth. Thus, according to one aspect of the present invention, the fusion protein is compartmentalized within cells of the plant or cultured plant cells, so as to be sequestered from cell walls of
10 the cells of the plant or cultured plant cells.

The fusion protein can be compartmentalized within a cellular compartment, such as, for example, the cytoplasm, endoplasmic reticulum, golgi apparatus, oil bodies, starch bodies, chloroplastids, chloroplasts, chromoplastids, chromoplasts, vacuole, lysosomes, mitochondria or the
15 nucleus.

Accordingly, the heterologous sequence used while implementing the process according to this aspect of the present invention includes (i) a first sequence encoding a cellulose binding peptide; (ii) a second sequence encoding a recombinant protein, wherein the first and second sequences are
20 joined together in frame; and (iii) a third sequence encoding a signal peptide for directing a protein to a cellular compartment, the third sequence being upstream and in frame with the first and second sequences.

The following provides description of signal peptides which can be used to direct the fusion protein according to the present invention to
25 specific cell compartments.

It is well-known that signal peptides serve the function of translocation of produced protein across the endoplasmic reticulum membrane. Similarly, transmembrane segments halt translocation and provide anchoring of the protein to the plasma membrane, see, Johnson *et al.* The Plant Cell (1990) 2:525-532; Sauer *et al.* EMBO J. (1990) 9:3045-3050; Mueckler *et al.* Science (1985) 229:941-945. Mitochondrial, nuclear, chloroplast, or vacuolar signals target expressed protein correctly into the corresponding organelle through the secretory pathway, see, Von Heijne, Eur. J. Biochem. (1983) 133:17-21; Von Heijne, J. Mol. Biol. (1986)
30 189:239-242; Iturriaga *et al.* The Plant Cell (1989) 1:381-390; McKnight *et al.*, Nucl. Acid Res. (1990) 18:4939-4943; Matsuoka and Nakamura, Proc. Natl. Acad. Sci. USA (1991) 88:834-838. A recent book by Cunningham and Porter (Recombinant proteins from plants, Eds. C. Cunningham and
35

In corn, within the kernel, proteins under the ubiquitin promoters, are preferentially accumulated in the germ (Kusnadi *et al.*, Biotechnol. Bioeng. (1998) 60:44-52). The amylose-extender (Ae) gene encoding starch-branching enzyme IIb (SBEIIb) in maize is predominantly expressed in endosperm and embryos during kernel development (Kim *et al.* Plant. Mol. Biol. (1998) 38:945-956). A starch branching enzyme (SBE) showed promoter activity after it was introduced into maize endosperm suspension cells by particle bombardment (Kim *et al.* Gene (1998) 216:233-243). In

transgenic wheat it has been shown that a native HMW-GS gene promoter can be used to obtain high levels of expression of seed storage and, potentially, other proteins in the endosperm (Blechl and Anderson, Nat. Biotechnol. (1996) 14:875-9). Polygalacturonase (PG) promoter was shown to confer high levels of ripening-specific gene expression in tomato (Nicholass *et al.* Plant. Mol. Biol. (1995) 28:423-435). The ACC oxidase promoter (Blume and Grierson, Plant. J. (1997) 12:731-746) represents a promoter from the ethylene pathway and shows increased expression during fruit ripening and senescence in tomato. The promoter for tomato 3-hydroxy-3-methylglutaryl coenzyme A reductase gene accumulates to high level during fruit ripening (Daraselia *et al.* Plant. Physiol. (1996) 112:727-733). Specific protein expression in potato tubers can be mediated by the patatin promoter (Sweetlove *et al.* Biochem. J. (1996) 320:487-492). Protein linked to a chloroplast transit peptide changed the protein content in transgenic soybean and canola seeds when expressed from a seed-specific promoter (Falco *et al.* Biotechnology (NY) (1995) 13:577-82). The seed specific bean phaseolin and soybean beta-conglycinin promoters are also suitable for the latter example (Keeler *et al.* Plant. Mol. Biol. (1997) 34:15-29). Promoters that are expressed in plastids are also suitable in conjunction with plastid transformation.

Each of these promoters can be used to implement the process according to the present invention.

Thus, the plant promoter employed can a constitutive promoter, a tissue specific promoter, an inducible promoter or a chimeric promoter.

Examples of constitutive plant promoters include, without being limited to, CaMV35S and CaMV19S promoters, FMV34S promoter, sugarcane bacilliform badnavirus promoter, CsVMV promoter, *Arabidopsis* ACT2/ACT8 actin promoter, *Arabidopsis* ubiquitin UBQ1 promoter, barley leaf thionin BTH6 promoter, and rice actin promoter.

Examples of tissue specific promoters include, without being limited to, bean phaseolin storage protein promoter, DLEC promoter, PHS β promoter, zein storage protein promoter, conglutin gamma promoter from soybean, AT2S1 gene promoter, ACT11 actin promoter from *Arabidopsis*, napA promoter from *Brassica napus* and potato patatin gene promoter.

The inducible promoter is a promoter induced by a specific stimuli such as stress conditions comprising, for example, light, temperature, chemicals, drought, high salinity, osmotic shock, oxidant conditions or in case of pathogenicity and include, without being limited to, the light-

inducible promoter derived from the pea *rbcS* gene, the promoter from the alfalfa *rbcS* gene, the promoters DRE, MYC and MYB active in drought; the promoters INT, INPS, *prxEa*, *Ha hsp17.7G4* and *RD21* active in high salinity and osmotic stress, and the promoters *hsr303J* and *str246C* active in pathogenic stress.

Expression follow-up:

Expression of the fusion protein can be monitored by a variety of methods. For example, ELISA or western blot analysis using antibodies specifically recognizing the recombinant protein or its cellulose binding peptide counterpart can be employed to qualitatively and/or quantitatively monitor the expression of the fusion protein in the plant. Alternatively, the fusion protein can be monitored by SDS-PAGE analysis using different staining techniques, such as, but not limited to, coomassie blue or silver staining. Other methods can be used to monitor the expression level of the RNA encoding for the fusion protein. Such methods include RNA hybridization methods, e.g., Northern blots and RNA dot blots.

Binding of the fusion protein to the plant derived cellulosic matter:

When sufficient expression has been detected, binding of the fusion protein to the plant derived cellulosic matter is effected. Such binding can be achieved, for example, as follows. Whole plants, plant derived tissue or cultured plant cells are homogenized by mechanical method in the presence or absence of a buffer, such as, but not limited to, PBS. The fusion protein is therefore given the opportunity to bind to the plant derived cellulosic matter. Buffers that may include salts and/or detergents at optimal concentrations may be used to wash non specific proteins from the cellulosic matter.

Extraction and purification:

In general, a recent book by Cunningham and Porter (Recombinant proteins from plants, Eds. C. Cunningham and A.J.R. Porter, 1998 Humana Press Totowa, N.J.) describes methods for the production of recombinant proteins in plants including methods for extraction of the proteins from the plants. The methods used herein for extraction of proteins from plants are similar, however the ability of the fusion protein to bind to cellulose dictates its fate, unless extraction is done under condition in which the cellulose binding peptide do not bind to cellulose, for example, pH higher than 10 (for most CBDs) or high concentration of glucose or cellobiose (200 mM or higher) for family IX CBDs. If the initial extraction is conducted under conditions that prevent binding, the supernatant is

cleared from the cellulosic matter and then the solution is brought by either dilution, dialysis or pH correction, if necessary, to a condition that enables binding, after which cellulose is added in a batch or the solution is loaded on a cellulose column. Cellulose affinity purification is conducted as described, for example, in U.S. Pat. Nos. 5,719,044; 5,670,623; 5,856,021; 5,137,819; 5,202,247; 5,340,731; and 5,474,925; and U.S. Pat. applications Nos. 08/788,621; and 08/788,622; EP 0 381 719 B1, and EP application No. 93907724.4. Alternatively, the extraction solution provides conditions that favor binding to the plant derived cellulosic matter.

In any case, while the fusion protein is bound to cellulose, further whases can be employed for further removal of unbound proteins, conditions which dissociate such binding or proteolytic cleavage can be used to isolate the fusion protein itself, or proteolytic cleavage can be used to isolate the recombinant protein, all as further detailed hereinabove.

Additional objects, advantages, and novel features of the present invention will become apparent to one ordinarily skilled in the art upon examination of the following examples, which is not intended to be limiting. Additionally, each of the various embodiments and aspects of the present invention as delineated hereinabove and as claimed in the claims section below finds experimental support in the following examples.

EXAMPLES

Reference is now made to the following examples, which together with the above descriptions, illustrate the invention in a non limiting fashion. Generally, the nomenclature used herein and the laboratory procedures in recombinant DNA technology described below are those well known and commonly employed in the art. Standard techniques are used for cloning, DNA and RNA isolation, amplification and purification. Generally enzymatic reactions involving DNA ligase, DNA polymerase, restriction endonucleases and the like are performed according to the manufacturers' specifications. These techniques and various other techniques are generally performed according to Sambrook *et al.*, molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989). Other general references are provided throughout this document. The procedures therein are believed to be well known in the art and are provided for the convenience of the reader. All the

information contained therein, as well as that contained in the Manual by Sambrook *et al.*, is incorporated herein by reference.

MATERIALS, CONSTRUCTS AND EXPERIMENTAL METHODS

5

Enzymes and Chemicals:

Chemicals were purchased from Sigma Israel Chemicals Ltd. (Rehovot, Israel) unless otherwise stated. Restriction enzymes were purchased from MBI Fermentas, Inc. (Amherst NY, USA) and Taq DNA
10 polymerase was purchased from Promega Corp. (Madison, WI, USA).

Plasmids and Bacteria:

The ligation mixture of each cloning procedure was transformed into *E. coli* strain XL1-blue (Stratagene) competent cells. The bacteria were plated on LB-agar plates including 100 µg/ml ampicillin in the case of the
15 pBlueScript and pUC plasmids, or 50 µg/ml kanamycin in the case of the shuttle vectors. Positive clones were verified by restriction analysis and sequencing.

Buffers and Media:

LB (Luria-Broth) - 1% bacto-tryptone, 0.5% yeast extract and 1%
20 NaCl; PBS (Phosphate Buffered Saline) - 20 mM KH₂PO₄, 150 mM NaCl pH 7.2; PBS-T - PBS with 0.1% Tween 20.

Plant Materials and Growth Conditions:

Nicotiana tabaccum-SR1 (tobacco) plants were grown at 24-25 °C under a 16 h photoperiod, using cool-white fluorescent light (50-60 µE m⁻²
25 S⁻¹).

Construction of ProtL-cex:

A vector containing the class-I patatin gene B33 promoter (Olesinski *et al.*, 1996, Plant Physiol. 111:541-550), fused to the cell signal sequence (Shani *et al.*, 1997, Plant Mol. Biol. 34:837-842), protein-L (hereinafter
30 ProtL, Nilson *et al.*, 1993, J. Immunol. Meth. 164:33-40), CBD cex sequence (Gilkes *et al.*, 1991, Microbiol. Rev. 55:303-315) and HDEL (SEQ ID NO:1) retaining peptide sequence was constructed as follows.

A DNA fragment encoding the cell signal peptide (nucleotide 1-105 of SEQ ID NO:3) was generated by PCR using the following primers: No. 1
35 (BglII) 5'-AAAAAGATCTATGGCGCGAAAATCACTAA-3' (SEQ ID NO:4) and No. 2 (XbaI) 5'-AAAATCTAGATTTACGGAGAGCGTCGCG-3' (SEQ ID NO:5). A DNA fragment encoding the ProtL-CBDcex (nucleotides 3-1280 of SEQ ID NO:6) was generated by PCR using the

following primers: No. 3 (XbaI) 5'-AAAATCTAGAATGGCG
GCGGTAGAAAATAAAG-3' (SEQ ID NO:7); and No. 4 (HDEL, Stop
and Sall) 5'-AAAAGTCGACTTAAAGTTCATCATGCTCGACGCC
GACCGTGCA-3' (SEQ ID NO:8). The two fragment were digested with
5 BglII, XbaI and Sall and ligated in one step into pUC19 (New England
Biolabs, Beverly, Massachusetts) that was pre digested with BamHI and
Sall. The primer for the c-terminal end of the ProtL-CBD contained the
retaining peptide sequence HDEL (SEQ ID NO:1) and a stop codon in
frame. The DNA containing the Cell Signal-ProtL-CBD-HDEL
10 (hereinafter, ProtL-CBD) fusion was excised using SmaI and Sall and was
subcloned into the SmaI and Sall sites of the binary vector Bin19 (Bevan,
1984, Nuc. Acid Res. 12:8711-8721) under the class-I patatin gene B33
promoter (Olesinski *et al.*, 1996, Plant Physiol. 111:541-550).

Construction of ProtL-cexNG (Non-Glycosylated):

15 The non-glycosylated form of CBDcex (CBD Technologies Ltd.) was
cloned into the Cell-ProtL-cex vector constructed as described above, thus
replacing the CBDcex with CBDcexNG. Cloning was performed using the
following PCR primers: Primer E, 5'-AAAAACTAGTGCTAGCGG
TCCAGCCGGC-3' (SEQ ID NO:9) which is a forward primer containing
20 an *SpeI* restriction site and primer F, 5'-AAAAGTCGACTTA
AAGTTCATCATGTCCAACGGTGCAAGGGGC-3' (SEQ ID NO:10)
which is a reverse primer containing the ER retaining peptide sequence
(HDEL), a stop codon and a *Sall* restriction site. The resultant 360 bp PCR
product was digested with *SpeI* and *Sall* and ligated into Cell-ProtL-cex
25 predigested with the same enzymes. Positive clones were sequenced for
verification and designated pUC19-cell-ProtL-cexNG-HDEL. Figure 2a
shows the coding sequence (SEQ ID NO:11) and the encoded protein (SEQ
ID NO:12) of construct pUC19-cell-ProtL-cexNG-HDEL, which is
schematically presented in Figure 2b.

30 ***Construction of the Shuttle Vector Containing a 35S- Ω Promoter:***

The Cell-ProtL-cexNG insert which was obtained by generated by
digesting the pUC19-cell-ProtL-cexNG-HDEL vector with *SmaI* and *SphI*
(*PaeI*) was sub cloned into a Cd vector containing the CaMV 35S- Ω
promoter. To ligate the Cell-ProtL-cexNG-HDEL insert, the Cd vector was
35 digested with *Sall* and the overhang tail was blunted by a fill-in reaction
using the Klenow fragment. The vector was then digested with *SphI*.

The resultant 35SΩ-cell1-ProtL-cexNG-HDEL vector and the pBI101 shuttle vector were digested via *Sma*I and *Sac*I and co-ligated to generate pBI-35SΩ-cell1-ProtL-cexNG-HDEL.

Construction of cex-Fx and Tma-Fx:

5 A CBDcex-Fx insert provided in a pBluescript II KS plasmid (pBS-cex-Fx) was obtained from Prof. Douglas Kilburn, Department of Microbiology and Immunology, Biotechnology Laboratory, The University of British Columbia, Vancouver. An ER retention peptide HDEL encoding sequence was ligated at the C-terminus of CBDcex-Fx as followed: Forward
10 primer 49, 5'-CTAGTCATGATGAACTTTAAGAGCT-3' (SEQ ID NO:13) and reverse primer 50, 5'-CTTAAAGTTCATCATGA-3' (SEQ ID NO:14) were mixed together at equi-molar ratios under denaturing conditions (94 ° C). The mixture was then allowed to cool to RT in order to allow annealing. The annealed primers were ligated into pBS-cex-Fx which was
15 predigested with *Spe*I and *Sac*I and the ligation mixture was used to transform XL1 blue competent cells. Positive clones were sequenced for verification and designated as pBS-cex-Fx-HDEL.

A Cell signal peptide encoding sequence was cloned into the N-terminus of pBS-cex-Fx-HDEL. The Cell signal peptide encoding
20 sequence was PCR amplified from pMH04 (Shani, Z., Dekel, M., Tsabary, G. and Shoseyov, O. (1997) Cloning and characterization of elongation specific endo-1,4-β-glucanase (cell) from *Arabidopsis thaliana*. Plant Molec. Biol. 34: 837-842.) using the following primers: Forward primer 51, 5'-AAAACCCGGGATGGCGCGAAAATC-3' (SEQ ID NO:15), containing
25 a *Sma*I restriction site, and reverse primer 52, 5'-AAAAGACGTCTTAC GGAGAGCGTCGCGGTAATC-3' (SEQ ID NO:16) containing an *Aat*II restriction site. The resulting 115 bp PCR product was digested with *Sma*I and *Aat*II and ligated into pBS-cex-Fx-HDEL. The ligation mixture was used to transform *E. coli* XL1 Blue competent cells. Positive clones were
30 verified via sequencing and designated as pBS-Sig-cex-Fx-HDEL. Figure 3a shows the coding sequence (SEQ ID NO:17) and the encoded protein (SEQ ID NO:18) of construct pBS-Sig-cex-Fx-HDEL, which is schematically presented in Figure 3b.

Replacing CBDcex with CBDTma:

35 CBDTma was PCR amplified from pET-CBDTma (Alam, M., Boraston, A.B., Kormos, J., Tomme, P. and Kilburn, D.G. Properties of the C-terminal family 9 cellulose-binding module of xylanase A from the hyperthermophilic bacterium *Thermatoga maritime*, Submitted) using

forward primer 53, 5'-AAAAGACGTCGGCTAGCGGAATAATGGTA
GCG-3', (SEQ ID NO:19), containing an *AatII* restriction site, and reverse
primer 54, 5'-AAAAACGCGTTGGGGATGGGGTCGGAC-3' (SEQ ID
NO:20), containing an *MluI* restriction site. The resultant 600 bp PCR
5 product was digested with *AatII* and *MluI* and ligated into pBS-Sig-cex-Fx-
HDEL that was predigested with the same enzymes. The ligation mixture
was used to transform *E.coli* XL1 Blue competent cells. Positive clones
were verified via sequencing and designated as pBS-Sig-Tma-Fx-HDEL.
Figure 4a shows the coding sequence (SEQ ID NO:21) and the encoded
10 protein (SEQ ID NO:22) of construct pBS-Sig-Tma-Fx-HDEL, which is
schematically presented in Figure 4b.

Shuttle Vector cloning:

A PJD-330 vector which contained the CaMV-35S- Ω promoter (a
kind donation from Prof. Gadi Galili, The Weizmann Institute, Rehovot,
15 Israel, Shaul, O. and Galili, G. (1992) Threonine overproduction in
transgenic tobacco plants expressing a mutant desensitized aspartate kinase
of *Escherichia coli*. Plant Physiol. 100: 1157-1163.) was digested with
HindIII and *SacI*. A 500 bp fragment was rescued and ligated into pBI101
(Clontech Laboratories Inc. Palo Alto, California, USA) predigested with
20 the same enzymes. The ligation mixture was used to transform *E.coli* XL1
Blue competent cells and verified positive clones were designated pBI-35S-
 Ω .

To clone CBDcex-Fx and CBDTma-Fx into the above described
shuttle vector, pBS-Sig-cex-Fx-HDEL and pBS-Sig-Tma-Fx-HDEL were
25 each digested with *SmaI* and *SacI* and respective 1.3 and 1.5 kb fragments
were rescued from these vectors and each ligated into pBI-35S- Ω
predigested with the same enzymes. The ligation mixture was used to
transform *E. coli* XL1 Blue competent cells and positive clones were
designated pBI-Sig-cex-Fx-HDEL and pBI-Sig-Tma-Fx-HDEL
30 respectively.

Plant Transformation:

The above described constructs were introduced into disarmed LB
4404 *Agrobacterium tumefaciens* by triparental mating (An, 1987, Meth.
Enzymol. 153: 292-305) and leaf-disc transformation was performed with
35 *Nicotiana tabacum*-SR1 plants as described previously (DeBlock et al.,
1984, EMBO J. 3:1681). Regenerated transgenic plants were selected on
kanamycin containing growth media and analyzed via PCR for the presence
of exogenic sequences as described below. Positive isolates were grown in

a tissue culture room or in a greenhouse and F₀ plants from independent transformation events were used for the protein purification assays.

Detection of Transgenic Plants by PCR from Chromosomal DNA:

DNA was extracted from leaves of *Nicotiana tabaccum*-SR1 (tobacco) as described by Doyle and Doyle (1987, *Phytochem. Bull.* 19:11-15) and the ProtL-cex sequence PCR amplified from the transgenic plant DNA isolated using the following primers: Forward primer 1, 5'-AAAACCATGGCGGCGGTAGAAAATAAAG-3' (SEQ ID NO:23) and reverse primer 2, 5'-AAAAGGATCCCTTCTGGTTTTTCGTCAAC -3' (SEQ ID NO:24).

In a similar manner, cex-Fx and Tma-Fx sequences were also PCR amplified from transgenic plant DNA by using the following primers: Forward primer 3, 5'-AAAACCCGGGATGGCGCGAAAATC-3' (SEQ ID NO:25) and reverse primer 4, 5'-TGC GTTCCAGGGTCTGTTTCC-3' (SEQ ID NO:26). The PCR reaction mixture included 2.5 µl 10X Taq polymerase buffer (Promega, Madison, WI), dNTP mix (0.2 mM each nucleotide), 1.5 mM MgCl₂, 10 pmol of each primer, 1 unit Taq DNA polymerase (Promega, Madison, WI) and ddH₂O to a final volume of 25 µl. Mineral oil (25 µl) was added to the mixture to prevent evaporation during cycling. The PCR program included 35 cycles of: denaturation at 95 °C for 60 sec, annealing at 55 °C - 65 °C for 1 min and extension at 72 °C for 1-2 min. The resulting amplified fragment was purified from an agarose-TBE gel.

Purification of ProtL-cexNG:

As outlined in Figure 5, purification of ProtL-cex was conducted as follow: 0.5 gram of tobacco leaf was grounded in liquid nitrogen. The grounded material was resuspended in PBS-T containing 2 mM PMSF, 5 mM EDTA and 2 mM DTT and the mixture was incubated at 4 °C for 1 hour with inversion to allow binding of the ProtL-cex to the cell wall fraction. The mixture was centrifuged, and the supernatant and pellet recovered separately. The pellet was washed 3 times with PBS-T and the washes were combined to the supernatant fraction which was then reacted with 10 mg Avicel 200. The washed cellulose pellet was washed 3 more times with PBS-T. The pellet and supernatant fractions were each separated on SDS-PAGE, and immunoblotted using either anti CBDcex Ab or mouse IgG reagent grades primary antibodies and appropriate secondary antibodies conjugated to HRP.

Purification of CBDcex-Fx and CBDTma-Fx:

Purification of CBDcex-Fx and CBDTma-Fx was conducted as described above. Analysis of the resultant pellet and supernatant fractions was conducted using the anti-Fx, anti-CBDcex or anti-CBDTma primary
5 antibodies and appropriate secondary antibodies conjugated to HRP.

EXPERIMENTAL RESULTS

Approximately 30 independent transgenic tobacco plants (F₀,
10 parental generation) were prepared from each of the ProtL-cex and Tma-Fx transformant lines. Confirmation of the presence of a transgene was conducted by kanamycin resistance and PCR analysis with specific primers as described above. The primers detected a 0.95 kb fragment in Prot-L transgenic plants (Figure 6), and 0.55 kb and 0.8 kb fragments from cex-Fx
15 (Figures 7a-7b) and Tma-Fx transgenic plants (Figures 8a-b), respectively. In all cases, the binary vector was used as a positive control. Expression of the cellulose binding domain (CBD) in the ProtL-cex and Tma-Fx transgenic plants was confirmed via western blot analysis.

Detection and Purification of ProtL-cexNG:

20 Of the positive transformants identified via PCR amplification, four plants (1, 2, 5 and 15) expressed ProtL-CBD to a detectable level. The total protein from leaf tissue of transformed plants was extracted and allowed to bind to the cell wall cellulose. The unbound protein in the soluble fraction of the total protein was allowed to bind to exogenous cellulose as described
25 in materials and methods. Western blot analysis of both fractions of the cell wall and the cellulose displayed a difference in the amount of ProtL-CBD present. ProtL-CBD was not detected in the cell wall fraction whereas in the cellulose fraction, a unique band was detected (Figure 9a). The detected ProtL-CBDcexNG was of a higher molecular weight (MW) as compared to
30 the bacterial ProtL-cex which was used as a positive control. This may be due to glycosilation of the plant expressed protein. The ability of the ProtL-cexNG to bind cellulose and the ability of protein L to bind mouse IgG even following gel analysis (Figure 9b), confirmed that the two bi-functional fusion proteins are active. The anti-CBDcex as well as the mouse IgG
35 western blots detected non-specific bands in the cell wall fraction. These non-specific bands appeared in the transgenic and wt plants, and represent a protein with a MW different than that expected for ProtL-cex.

By correlating the brightness of the signal specific band observed on the western blot with bands of ProtL-CBD isolated from *E.coli*, the amount of ProtL-CBDcexNG accumulated in the transformed plant tissue was approximated to be 1 µg fusion protein per gram of plant tissue.

5 ***Detection and Purification of cex-Fx:***

Cex-Fx transformants were examined for their ability to express cex-Fx. Two transformant lines (5 and 12) expressed the protein to a detectable level. CBDcex-Fx was detected in the cell wall and cellulose fractions (prepared as described above) in equal amounts (Figures 10a-b).
10 The amount of cex-Fx produced in the plant tissue was approximated at 5 µg fusion protein per gram of plant tissue. As is evident from Figures 10a-b, the expressed fusion protein is of a higher MW then control CBDcex-Fxa which is expressed in mammalian cells. This shift in MW could be a result of inefficient processing of the protein, at the kex2 and Fx cleavage sites.

15 ***Detection and Purification of Tma-Fx:***

Tma-Fx transformants were examined for their ability to express Tma-Fx. Four transformants (11, 14, 17 and 19) expressed the protein to a detectable level. The expressed CBDTma-Fx was found only in the exogenous cellulose fraction (prepared as described above) indicating that
20 the CBDTma did not bind cell wall in the transformant plants (Figures 11a-d). This could be due to the high concentration of endogenous soluble sugars such as glucose, and cellobiose, which prevent the CBDTma from binding to the cellulose matrix. The soluble sugars in the supernatant fraction are diluted by the repeated washes and as such, the CBDTma
25 accumulated in this fraction is able to bind with exogenously added cellulose.

Further support for this theory can be found in the results of transformant 19 (Figures 11c-d). The amount of tissue recovered from this plant for extraction was significantly lower (4-10 times) then that recovered
30 from the other plants. Since the final volume of the samples was equal, the sugar concentration in the sample extracted from transformant 19 was lower, resulting in binding to the cell wall fraction.

The expressed fusion protein appeared to be of a higher MW then CBDcex-Fxa expressed in mammalian cells (positive control). This shift in
35 MW could be a result of inefficient processing of the protein, at the kex2 and Fx cleavage sites. The amount of Tma-fx accumulated in the transformant plant tissue was approximately 5 µg of fusion protein per gram of plant tissue.

Thus, as clearly shown by the preceding examples, expression of and cell wall isolation of exogenous proteins in plant tissue can easily be facilitated by utilizing any of the cellulose binding peptides of the present invention.

5

Although the invention has been described in conjunction with specific embodiments thereof, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, it is intended to embrace all such alternatives, modifications and variations that fall within the spirit and broad scope of the appended claims. All publications, patents and patent applications mentioned in this specification are herein incorporated in their entirety by reference into the specification, to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated herein by reference. In addition, citation or identification of any reference in this application shall not be construed as an admission that such reference is available as prior art to the present invention.

10

15

WHAT IS CLAIMED IS:

1. A process of expressing a recombinant protein in a plant and of isolating the recombinant protein from the plant, the process comprising the steps of:

- (a) providing a plant, a plant derived tissue or cultured plant cells expressing a fusion protein including the recombinant protein and a cellulose binding peptide being fused thereto, said fusion protein being compartmentalized within cells of said plant, plant derived tissue or cultured plant cells, so as to be sequestered from cell walls of said cells of said plant, plant derived tissue or cultured plant cells;
- (b) homogenizing said plant, plant derived tissue or cultured plant cells, so as to bring into contact said fusion protein with a plant derived cellulosic matter of said plant, plant derived tissue or cultured plant cells, to thereby effect affinity binding of said fusion protein via said cellulose binding peptide to said cellulosic matter, thereby obtaining a fusion protein cellulosic matter complex; and
- (c) isolating said fusion protein cellulosic matter complex.

2. The process of claim 1, further comprising the step of:

- (d) washing said fusion protein cellulosic matter complex, thereby removing endogenous plant proteins and other plant material therefrom.

3. The process of claim 2, further comprising the step of:

- (e) collecting said fusion protein cellulosic matter complex as a final product of the process.

4. The process of claim 2, further comprising the step of:

- (e) exposing said fusion protein cellulosic matter complex to conditions effective in dissociating said fusion protein from said cellulosic matter; and
- (f) isolating said fusion protein, thereby obtaining an isolated fusion protein.

5. The process of claim 4, wherein said conditions effective in dissociating said fusion protein from said cellulosic matter are selected from the group consisting of basic conditions, denaturative conditions and affinity displacement conditions.

6. The process of claim 4, further comprising the step of:
- (g) exposing said isolated fusion protein to conditions effective in digesting said fusion protein so as to release said recombinant protein therefrom, thereby obtaining a released recombinant protein.

7. The process of claim 6, wherein said conditions effective in digesting said fusion protein so as to release said recombinant protein therefrom are selected from the group consisting of proteolysis effected via a protease and proteolysis effected under predetermined cis or trans conditions for digesting a controllable intervening protein sequence.

8. The process of claim 4, further comprising the step of:
- (h) isolating said released recombinant protein.

9. The process of claim 2, further comprising the step of:
- (e) exposing said fusion protein cellulosic matter complex to conditions effective in digesting said fusion protein so as to release said recombinant protein therefrom, thereby obtaining a released recombinant protein.

10. The process of claim 6, wherein said conditions effective in digesting said fusion protein so as to release said recombinant protein therefrom are selected from the group consisting of proteolysis effected via a protease and proteolysis effected under predetermined cis or trans conditions for digesting a controllable intervening protein sequence.

11. The process of claim 4, further comprising the step of:
- (f) isolating said released recombinant protein.

12. A genetically modified or viral infected plant or cultured plant cells expressing a fusion protein including a recombinant protein and a cellulose binding peptide, said fusion protein being compartmentalized

within cells of said plant or cultured plant cells, so as to be sequestered from cell walls of said cells of said plant or cultured plant cells.

13. The genetically modified or viral infected plant or cultured plant cells of claim 12, wherein expression of said fusion protein is under a control of a constitutive or tissue specific plant promoter.

14. The genetically modified or viral infected plant or cultured plant cells of claim 12, wherein said fusion protein is compartmentalized within a cellular compartment selected from the group consisting of cytoplasm, endoplasmic reticulum, golgi apparatus, oil bodies, starch bodies, chloroplastids, chloroplasts, chromoplastids, chromoplasts, vacuole, lysosomes, mitochondria, and nucleus.

15. A genetically modified or viral infected plant or cultured plant cells expressing a fusion protein including a recombinant protein and a cellulose binding peptide separated therebetween via a unique amino acid sequence recognizable and digestible by a protease or under predetermined cis or trans conditions for digesting a controllable intervening protein sequence.

16. The genetically modified or viral infected plant or cultured plant cells of claim 15, wherein said fusion protein is compartmentalized within cells of said plant or cultured plant cells, so as to be sequestered from cell walls of said cells of said plant or cultured plant cells.

17. The genetically modified or viral infected plant or cultured plant cells of claim 15, wherein said fusion protein is compartmentalized within a cellular compartment selected from the group consisting of cytoplasm, endoplasmic reticulum, golgi apparatus, oil bodies, starch bodies, chloroplastids, chloroplasts, chromoplastids, chromoplasts, vacuole, lysosomes, mitochondria, and nucleus.

18. The genetically modified or viral infected plant or cultured plant cells of claim 15, wherein expression of said fusion protein is under a control of a constitutive or tissue specific plant promoter.

19. A composition of matter comprising:
- (a) a plant derived cellulosic matter of a plant; and
 - (b) a fusion protein including a recombinant protein and a cellulose binding peptide separated therebetween via a unique amino acid sequence recognizable and digestible by a protease or under predetermined cis or trans conditions for digesting a controllable intervening protein sequence, said fusion protein being expressed in said plant and complexed to said plant derived cellulosic matter of said plant by affinity binding via said cellulose binding peptide.
20. A nucleic acid molecule comprising:
- (a) a promoter sequence for directing protein expression in plant cells;
 - (b) a heterologous nucleic acid sequence including:
 - (i) a first sequence encoding a cellulose binding peptide;
 - (ii) a second sequence encoding a recombinant protein, wherein said first and second sequences are joined together in frame; and
 - (iii) a third sequence encoding a unique amino acid sequence being recognizable and digestible by a protease or under predetermined cis or trans conditions for digesting a controllable intervening protein sequence, said third sequence being between and in frame with said first and second sequences;

wherein, said heterologous nucleic acid sequence being down stream said promoter sequence, such that expression of said heterologous nucleic acid sequence is effectable by said promoter sequence.

21. The nucleic acid molecule of claim 20, further comprising a sequence element selected from the group consisting of an origin of replication for propagation in bacterial cells, at least one sequence element for integration into a plant's genome, a polyadenylation recognition sequence, a transcription termination signal, a sequence encoding a translation start site, a sequence encoding a translation stop site, plant RNA virus derived sequences, plant DNA virus derived sequences, tumor inducing (Ti) plasmid derived sequences, and a transposable element derived sequence.

22. A nucleic acid molecule comprising:
- (a) a promoter sequence for directing protein expression in plant cells;
 - (b) a heterologous nucleic acid sequence including:
 - (i) a first sequence encoding a cellulose binding peptide;
 - (ii) a second sequence encoding a recombinant protein, wherein said first and second sequences are joined together in frame; and
 - (iii) a third sequence encoding a signal peptide for directing a protein to a cellular compartment, said third sequence being upstream and in frame with said first and second sequences;

wherein, said heterologous nucleic acid sequence being down stream said promoter sequence, such that expression of said heterologous nucleic acid sequence is effectable by said promoter sequence.

23. The nucleic acid molecule of claim 22, wherein said heterologous nucleic acid sequence further includes

- (iv) a fourth sequence encoding a unique amino acid sequence being recognizable and digestible by a protease or under predetermined cis or trans conditions for digesting a controllable intervening protein sequence, said fourth sequence being between and in frame with said first and second sequences.

24. The nucleic acid molecule of claim 22, further comprising a sequence element selected from the group consisting of an origin of replication for propagation in bacterial cells, at least one sequence element for integration into a plant's genome, a polyadenylation recognition sequence, a transcription termination signal, a sequence encoding a translation start site, a sequence encoding a translation stop site, plant RNA virus derived sequences, plant DNA virus derived sequences, tumor inducing (Ti) plasmid derived sequences, and a transposable element derived sequence.

25. The nucleic acid molecule of claim 22, wherein said cellular compartment is selected from the group consisting of cytoplasm, endoplasmic reticulum, golgi apparatus, oil bodies, starch bodies,

chloroplastids, chloroplasts, chromoplastids, chromoplasts, vacuole, lysosomes, mitochondria, and nucleus.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 December 2000 (21.12.2000)

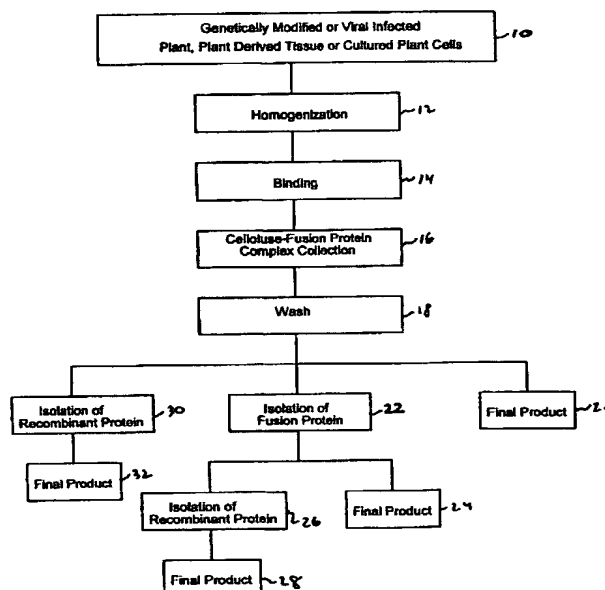
PCT

(10) International Publication Number
WO 00/77174 A1

- (51) International Patent Classification⁷: C12N 5/04, 15/09, 15/62, 15/64, 15/67, 15/82, 15/90, A01H 5/00
- (21) International Application Number: PCT/IL00/00330
- (22) International Filing Date: 7 June 2000 (07.06.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 09/329,234 10 June 1999 (10.06.1999) US
- (71) Applicants (for all designated States except US): CBD TECHNOLOGIES LTD. [IL/IL]; 4th Floor, 2 Pekeris, Park Tamar, 76100 Rehovot (IL). YISSUM RESEARCH DEVELOPMENT COMPANY OF THE HEBREW UNIVERSITY OF JERUSALEM [IL/IL]; 46 Jabotinsky Street, 91042 Jerusalem (IL).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): SHANI, Ziv [IL/IL]; 4 Paldi Street, 76248 Rehovot (IL). SHOSEYOV, Oded [IL/IL]; 5 Haerez Street, 72910 Karne Yosef (IL).
- (74) Agent: G.E. EHRLICH (1995) LTD.; 17th Floor, 28 Bezael Street, 52521 Ramat Gan (IL).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European

[Continued on next page]

(54) Title: PROCESS OF EXPRESSING AND ISOLATING RECOMBINANT PROTEINS AND RECOMBINANT PROTEIN PRODUCTS FROM PLANTS, PLANT DERIVED TISSUES OR CULTURED PLANT CELLS



(57) Abstract: A process of expressing a recombinant protein in a plant and of isolating the recombinant protein from the plant, the process is effected by (a) providing a plant, a plant derived tissue or cultured plant cells expressing a fusion protein including the recombinant protein and a cellulose binding peptide being fused thereto, the fusion protein being compartmentalized within cells of the plant, plant derived tissue or cultured plant cells, so as to be sequestered from cell walls of the cells of the plant, plant derived tissue or

[Continued on next page]

WO 00/77174 A1

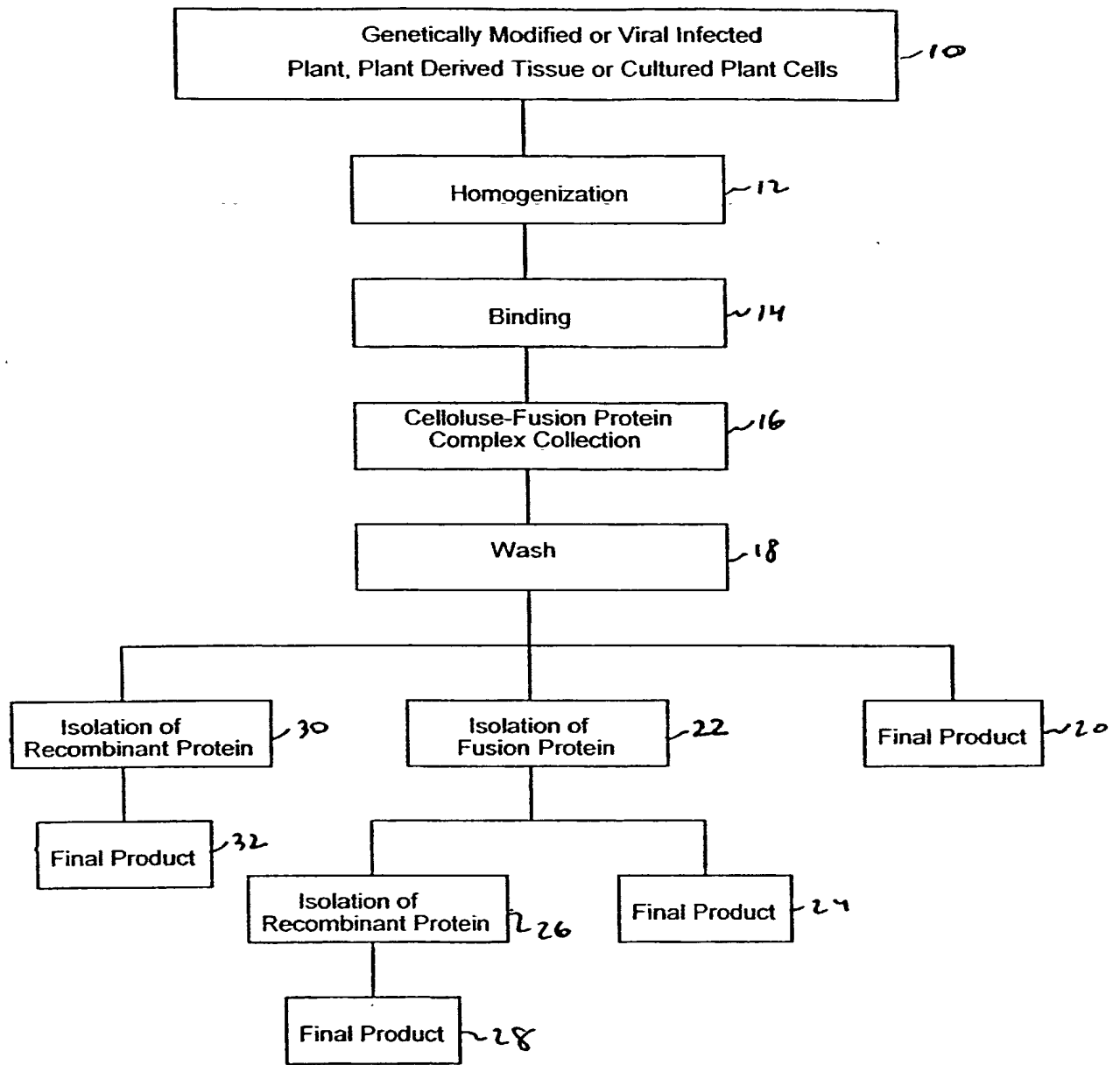


FIG. 1

1 GGGGATCTATGGCGCGAAAATCCCTAATTTTCCCGGTGATTTTGCTCGCCGTTCTTCTCT
1 M A R K S L I F P V I L L A V L L

61 TCTCTCCGCCGATTACTCCGCCGGTCACGATTACCGCGACGCTCTCCGTAAATCTAGAA
18 F S P P I Y S A G H D Y R D A L R K S R

121 TGGCGGCGGTAGAAAATAAAGAAGAAACACCAGAAACACCAGAAACTGATTCAAGAAG
38 M A A V E N K E E T P E T P E T D S E E

181 AAGTAACAATCAAAGCTAACCTAATCTTTGCAAATGGAAGCACACAACTGCAGAATTCA
58 E V T I K A N L I F A N G S T Q T A E F

241 AAGGAACATTTGAAAAAGCAACATCAGAAGCTTATGAGTATGCAGATACTTTGAAGAAAG
78 K G T F E K A T S E A Y E Y A D T L K K

301 ACAATGGAGAATATACTGTAGATGTTGCAGATAAAGGTTATACTTTAAATATTAAATTTG
98 D N G E Y T V D V A D K G Y T L N I K F

361 CTGGAAAAGAAAAACACCAGAAGAACCAAAAGAAGAAGTTACTATTAAAGCAAACCTAA
118 A G K E K T P E E P K E E V T I K A N L

421 TCTATGCAGATGGAAAAACACAAACAGCAGAATTCAAAGGAACATTTGAAGAAGCAACAG
138 I Y A D G K T Q T A E F K G T F E E A T

481 CAGAAGCATACAGATATGCAGATGCATTAAAGAAGGACAATGGAGAATATACAGTAGACG
158 A E A Y R Y A D A L K K D N G E Y T V D

541 TTGCAGATAAAGGTTATACTTTAAATATTAAATTTGCTGGAAAAGAAAAACACCAGAAG
178 V A D K G Y T L N I K F A G K E K T P E

601 AACCAAAAGAAGAAGTTACTATTAAAGCAAACCTTAATCTATGCAGATGGAAAAACACAAA
198 E P K E E V T I K A N L I Y A D G K T Q

661 CAGCAGAATTCAAAGGAACATTTGAAGAAGCAACAGCAGAAGCATACAGATATGCTGACT
218 T A E F K G T F E E A T A E A Y R Y A D

721 TATTAGCAGCAAAAGAAAATGGTAAATATACAGTAGACGTTGCAGATAAAGGTTATACTT
238 L L A A K E N G K Y T V D V A D K G Y T

781 TAAATATTAAATTTGCTGGAAAAGAAAAACACCAGAAGAACCAAAAGAAGAAGTTACTA
258 L N I K F A G K E K T P E E P K E E V T

841 TTAAAGCAAACCTTAATCTATGCAGATGGAAAACTCAAACAGCAGAGTTCAAAGGAACAT
278 I K A N L I Y A D G K T Q T A E F K G T

901 TTGCAGAAGCAACAGCAGAAGCATACAGATACGCTGACTTATTAGCAAAAGAAAATGGTA
298 F A E A T A E A Y R Y A D L L A K E N G

961 AATATACAGCAGACTTAGAAGATGGTGGATACACTATTAATATTAGATTTGCAGGTAAGA
318 K Y T A D L E D G G Y T I N I R F A G K

1021 AAGTTGACGAAAAACCAGAAGGGATCCCTCCGACGCCGACCCCGACTAGTGCTAGCGGTC
338 K V D E K P E G I P P T P T P T S A S G

Fig. 2a

Fig. 2a CONT.

1081	CAGCCGGCTGCCAGGTTCTGTGGGGTGTAAACCAGTGGAACACCGGTTTCACCGCTCAGG
358	P A G C Q V L W G V N Q W N T G F T A Q
1141	TTACCGTTAAAAACACGGGCTCAGCTCCGGTTGACGGTTGGACCCTGACCTTCTCTTTTC
378	V T V K N T G S A P V D G W T L T F S F
1201	CCTCGGGTCAGCAGGTAACCTCAGGCTTGGTCATCTACAGTTACCCAGTCTGGATCCGCTG
398	P S G Q Q V T Q A W S S T V T Q S G S A
1261	TTACAGTTCGTAAACGCTCCGTGGAACGGTAATATTCCTGCAGGTGGAACCGCTCAGTTCCG
418	V T V R N A P W N G N I P A G G T A Q F
1321	GTTTCCAAGGTTCTCACACCGGTACCAACGCGGCCCAACCGCTTCTCTCTGAACGGGG
438	G F Q G S H T G T N A A P T A F S L N G
1381	CCCCTTGCACCGTTGGACATGATGAACTTTAAGTCGAC
458	A P C T V G H D E L

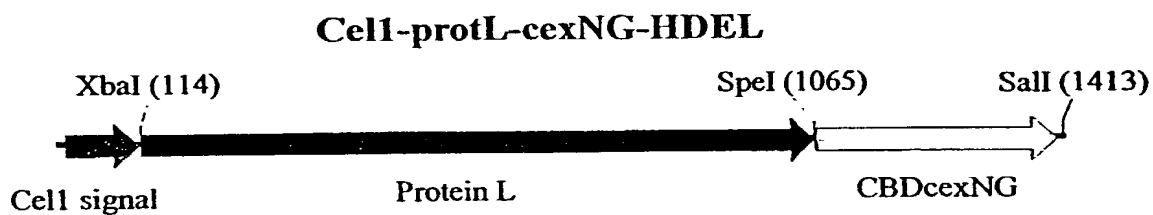


Fig. 2b

61 G C C C C C C C T C G A G C C C G G G A T G G C G C G A A A A T C C C T A A T T T T C C C G G T G A T T T T G C T C G C
 1 M A R K S L I F P V I L L A
 121 C G T T C T T C T T C T C T C C G C C G A T T T A C T C C G C C G G T C A C G A T T A C C G C G A C G C T C T C C G
 15 V L L F S P P I Y S A G H D Y R D A L R
 181 T A A G A C G T C C G G T C C A G C C G G C T G C C A G G T T C T G T G G G G T G T T A A C C A G T G G A A C A C C G G
 35 K T S G P A G C Q V L W G V N Q W N T G
 241 T T T C A C C G C T C A G G T T A C C G T T A A A A A C A C G G G C T C A G C T C C G G T T G A C G G T T G G A C C C T
 55 F T A Q V T V K N T G S A P V D G W T L
 301 G A C C T T C T C T T T T C C C T C G G G T C A G C A G G T A A C T C A G G C T T G G T C A T C T A C A G T T A C C C A
 75 T F S F P S G Q Q V T Q A W S S T V T Q
 361 G T C T G G A T C C G C T G T T A C A G T T C G T A A C G C T C C G T G G A A C G G T A A T A T T C C T G C A G G T G G
 95 S G S A V T V R N A P W N G N I P A G G
 421 A A C C G C T C A G T T C G G T T T C C A A G G T T C T C A C A C C G G T A C C A A C G C G G C C C A A C C G C T T T
 115 T A Q F G F Q G S H T G T N A A P T A F
 481 C T C T C T G A A C G G G G C C C C T T G C A C C G T T G G T C C G A C T A C C T C A C T A C A A C G C G T A A G C T
 135 S L N G A P C T V G P T T S P T T R K L
 541 C T G C A G C C T G G A C A A C G G G A C T G T G A C C A G T T C T G C C A C G A G G A A C A G A A C T C T G T G G T
 155 C S L D N G D C D Q F C H E E Q N S V V
 601 G T G C T C C T G C G C C C G C G G T A C A C C T G G C T G A C A A C G G C A A G G C C T G C A T T C C C A C A G G
 175 C S C A R G Y T L A D N G K A C I P T G
 661 G C C C T A C C C C T G T G G G A A C A G A C C C T G G A A C G C A G G A A G A G G T C A G T G G C C C A G G C C A C
 195 P Y P C G K Q T L E R R K R S V A Q A T
 721 C A G C A G C A G C G G G A G G C C C C T G A C A G C A T C A C A T G G A A G C C A T A T G A T G C A G C C G A C C T
 215 S S S G E A P D S I T W K P Y D A A D L
 781 G G A C C C C A C C G A G A A C C C C T T C G A C C T G C T T G A C T T T G A T C A G A C G C A G C C T G A G A G G G G
 235 D P T E N P F D L L D F D Q T Q P E R G
 841 C G A C A A C A A C A T T G A A G G T C G T A T C G T G G G A G G C C A G G A A T G C A A G G A C G G G G A G T G T C C
 255 D N N I E G R I V G G Q E C K D G E C P
 901 C T G G C A G G C C C T G C T C A T C A A T G A G G A A A C G A G G G T T C T G T G G T G G A A C C A T T C T G A G
 275 W Q A L L I N E E N E G F C G G T I L S
 961 C G A G T T C T A C A T C C T A A C G G C A G C C C A C T G T C T C T A C C A A G C C A A G A G A T T C A A G G T G A G
 295 E F Y I L T A A H C L Y Q A K R F K V R
 1021 G G T A G G G A C C G G A A C A C G G A G C A G G A G G G C G G T G A G G C G G T G C A C G A G G T G G A G G T
 315 V G D R N T E Q E E G G E A V H E V E V
 1081 G G T C A T C A A G C A C A A C C G G T T C A C A A G G A G A C C T A T G A C T T C G A C A T C G C C G T G C T C C G
 335 V I K H N R F T K E T Y D F D I A V L R
 1141 G C T C A A G A C C C C C A T C A C C T T C C G C A T G A A C G T G G C G C C T G C C T G C C T C C C C G A G C G T G A
 355 L K T P I T F R M N V A P A C L P E R D

Fig. 3a

FIG. 3a CONT.

1201	CTGGGCCGAGTCCACGCTGATGACGCAGAAGACGGGGATTGTGAGCGGCTTCGGGCGCAC
375	W A E S T L M T Q K T G I V S G F G R T
1261	CCACGAGAAGGGCCGGCAGTCCACCAGGCTCAAGATGCTGGAGGTGCCCTACGTGGACCG
395	H E K G R Q S T R L K M L E V P Y V D R
1321	CAACAGCTGCAAGCTGTCCAGCAGCTTCATCATCACCCAGAACATGTTCTGTGCCGGCTA
415	N S C K L S S S F I I T Q N M F C A G Y
1381	CGACACCAAGCAGGAGGATGCCTGCCAGGGGGACAGCGGGGGCCCGCACGTCACCCGCTT
435	D T K Q E D A C Q G D S G G P H V T R F
1441	CAAGGACACCTACTTCGTGACAGGCATCGTCAGCTGGGGAGAGGGCTGTGCCCGTAAGGG
455	K D T Y F V T G I V S W G E G C A R K G
1501	GAAGTACGGGATCTACACCAAGGTCACCGCCTTCCTCAAGTGGATCGACAGGTCCATGAA
475	K Y G I Y T K V T A F L K W I D R S M K
1561	AACCAGGGGCTTGCCCAAGCCAAGCCTACTAGTCATGATGAACCTTTAAGAGCTCCAGCT
495	T R G L P K A K P T S H D E L

Sig-cex-Fx-HDEL

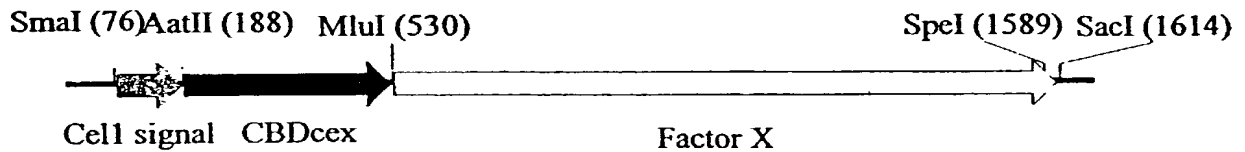


FIG. 3b

61 G C C C C C C T C G A G C C C G G G A T G G C G C G A A A A T C C C T A A T T T T C C C G G T G A T T T T G C T C G C
1 M A R K S L I F P V I L L A

121 C G T T C T T C T T C T C T C C G C C G A T T T A C T C C G C C G G T C A C G A T T A C C G C G A C G C T C T C C G
15 V L L F S P P I Y S A G H D Y R D A L R

181 T A A G A C G T C G G C T A G C G G A A T A T G G T A G C G A C A G C A A A A T A C G G A A C A C C G G T C A T C G A
35 K T S A S G I M V A T A K Y G T P V I D

241 T G G A G A G A T A G A C G A G A T C T G G A A C A C G A C A G A G G A G A T A G A G A C G A A A G C G G T G G C C A T
55 G E I D E I W N T T E E I E T K A V A M

301 G G G A T C G C T T G A C A A G A A C G C G A C A G C G A A A G T G A G G G T G C T G T G G G A C G A G A A C T A C C T
75 G S L D K N A T A K V R V L W D E N Y L

361 G T A C G T A C T T G C A A T C G T G A A A G A C C C C G T T C T G A A C A A G A C A A C A G C A A C C C G T G G G A
95 Y V L A I V K D P V L N K D N S N P W E

421 A C A G G A T T C C G T G G A G A T C T T C A T C G A C G A G A A C A A C C A A G A C A G G A T A C T A C G A A G A
115 Q D S V E I F I D E N N H K T G Y Y E D

481 C G A C G A C G C A C A G T T C A G G G T G A A C T A C A T G A A C G A G C A G A C G T T T G G A A C G G G A G G A A G
135 D D A Q F R V N Y M N E Q T F G T G G S

541 T C C A G C G A G G T T C A A G A C A G C G G T G A A A C T G A T C G A A G G A G G A T A C A T A G T T G A G G C A G C
155 P A R F K T A V K L I E G G Y I V E A A

601 G A T C A A G T G G A A G A C G A T C A A A C C C A C A C C G A A C A C G G T G A T A G G A T T C A A C A T C C A G G T
175 I K W K T I K P T P N T V I G F N I Q V

661 G A A C G A T G C G A A C G A G A A A G G G C A G A G G G T C G G T A T C A T C T C C T G G A G C G A T C C C A C A A A
195 N D A N E K G Q R V G I I S W S D P T N

721 C A A C A G C T G G A G A G A T C C T T C A A A G T T C G G T A A C C T C A G G C T C A T C A A G G G A T C T G G T C C
215 N S W R D P S K F G N L R L I K G S G P

781 G A C C C C A T C C C C A A C G C G T A A G C T C T G C A G C C T G G A C A A C G G G G A C T G T G A C C A G T T C T G
235 T P S P T R K L C S L D N G D C D Q F C

841 C C A C G A G G A A C A G A A C T C T G T G G T G T G C T C C T G C G C C C G C G G T A C A C C C T G G C T G A C A A
255 H E E Q N S V V C S C A R G Y T L A D N

901 C G G C A A G G C C T G C A T T C C C A C A G G G C C C T A C C C C T G T G G G A A C A G A C C C T G G A A C G C A G
275 G K A C I P T G P Y P C G K Q T L E R R

961 G A A G A G G T C A G T G G C C C A G G C C A C C A G C A G C A G C G G G G A G G C C C C T G A C A G C A T C A C A T G
295 K R S V A Q A T S S S G E A P D S I T W

1021 G A A G C C A T A T G A T G C A G C C G A C C T G G A C C C C A C C G A G A A C C C C T T C G A C C T G C T T G A C T T
315 K P Y D A A D L D P T E N P F D L L D F

1081 T G A T C A G A C G C A G C C T G A G A G G G G C A C A A C A A C A T T G A A G G T C G T A T C G T G G G A G G C C A
335 D Q T Q P E R G D N N I E G R I V G G Q

1141 G G A A T G C A A G G A C G G G G A G T G T C C C T G G C A G G C C C T G C T C A T C A A T G A G G A A A C G A G G G
355 E C K D G E C P W Q A L L I N E E N E G

Fig. 4a

Fig. 4a cont.

1201 TTTCTGTGGTGGAAACCATTCTGAGCGAGTTCTACATCCTAACGGCAGCCCACTGTCTCTA
 375 F C G G T I L S E F Y I L T A A H C L Y

 1261 CCAAGCCAAGAGATTCAAGGTGAGGGTAGGGGACCGGAACACGGAGCAGGAGGAGGGCGG
 395 Q A K R F K V R V G D R N T E Q E E G G

 1321 TGAGCGGTGCACGAGGTGGAGGTGGTCATCAAGCACAACCGGTTACAAAAGGAGACCTA
 415 E A V H E V E V V I K H N R F T K E T Y

 1381 TGACTTCGACATCGCCGTGCTCCGGCTCAAGACCCCCATCACCTCCGCATGAACGTGGC
 435 D F D I A V L R L K T P I T F R M N V A

 1441 GCCTGCCTGCCTCCCCGAGCGTGACTGGGCCGAGTCCACGCTGATGACGCAGAAGACGGG
 455 P A C L P E R D W A E S T L M T Q K T G

 1501 GATTGTGAGCGGCTTCGGGCGCACCCACGAGAAGGGCCGGCAGTCCACCAGGCTCAAGAT
 475 I V S G F G R T H E K G R Q S T R L K M

 1561 GCTGGAGGTGCCCTACGTGGACCGCAACAGCTGCAAGCTGTCCAGCAGCTTCATCATCAC
 495 L E V P Y V D R N S C K L S S S F I I T

 1621 CCAGAACATGTTCTGTGCCGGCTACGACACCAAGCAGGAGGATGCCAGGGGGACAG
 515 Q N M F C A G Y D T K Q E D A C Q G D S

 1681 CGGGGGCCCGCACGTCACCCGCTTCAAGGACACCTACTTCGTGACAGGCATCGTCAGCTG
 535 G G P H V T R F K D T Y F V T G I V S W

 1741 GGGAGAGGGCTGTGCCCGTAAGGGGAAGTACGGGATCTACACCAAGGTCACCGCCTTCT
 555 G E G C A R K G K Y G I Y T K V T A F L

 1801 CAAGTGGATCGACAGGTCCATGAAAACCAGGGGCTTGCCCAAGGCCAAGCCTACTAGTCA
 575 K W I D R S M K T R G L P K A K P T S H

 1861 TGATGAACTTTAAGAGCTCCAGCTTTTGTTCCTTTAGTGAGGGTTAATTGCGCGCTTGG
 595 D E L

Cell-TmaIX-Fx-HDEL

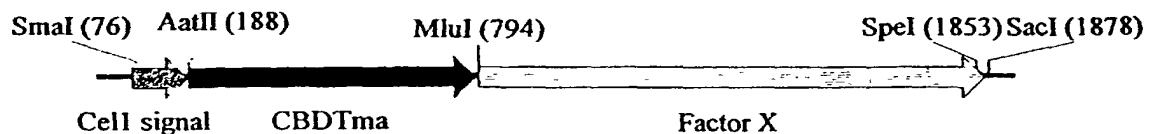


Fig. 4b

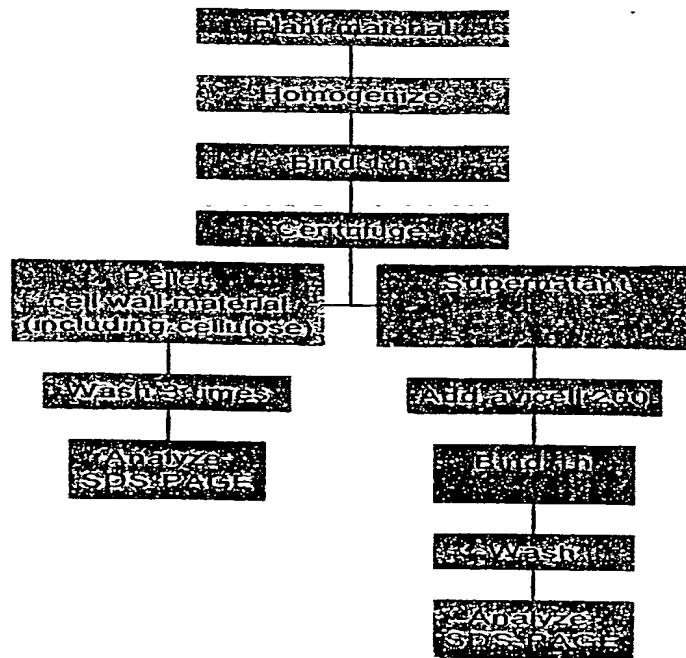


Fig. 5

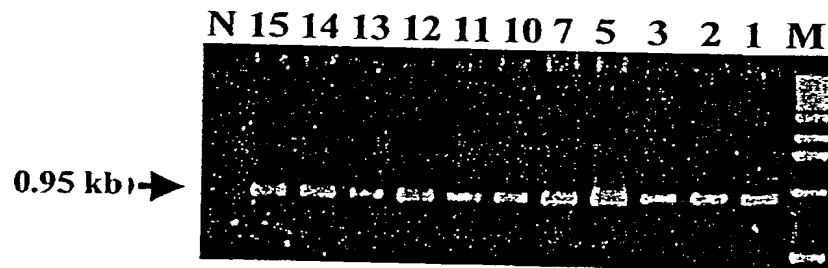


Fig. 6

13 12 11 10 9 8 7 6 5 4 3 2 1 M

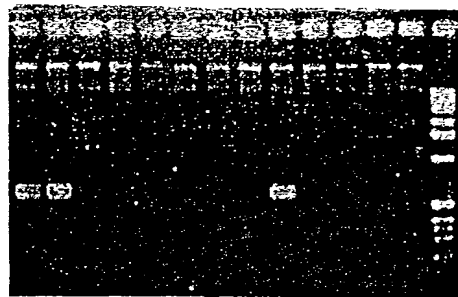


FIG. 7a

N19 18 17 16 M

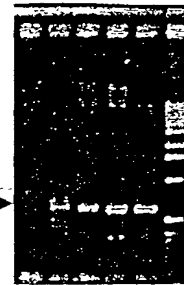


FIG. 7b

11 10 9 8 7 6 5 4 3 2 1 M

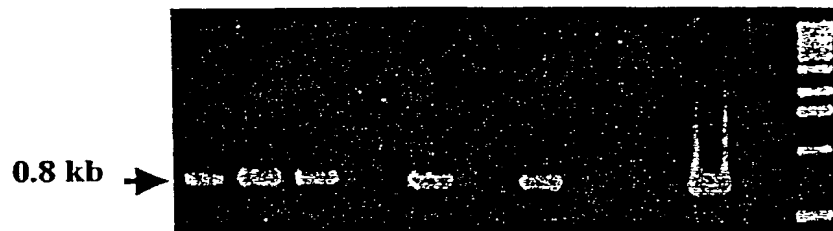


FIG. 8a

C 22 21 20 19 18 17 16 15 14 13 12 M

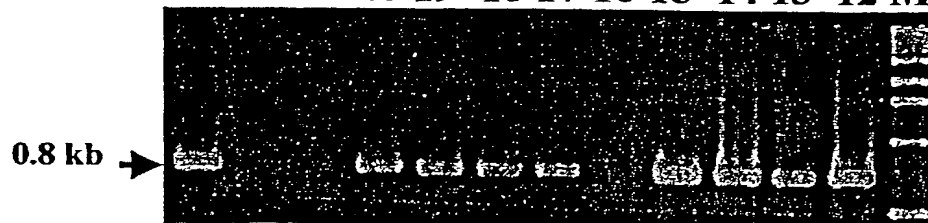


FIG. 8b

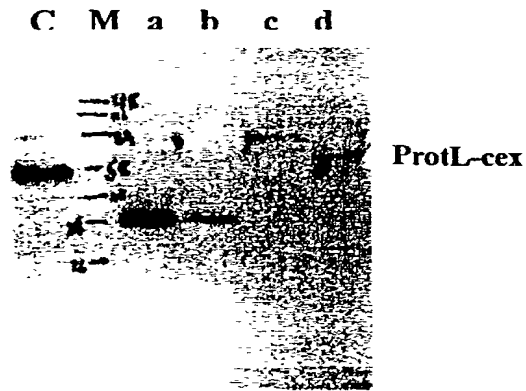


FIG. 9a

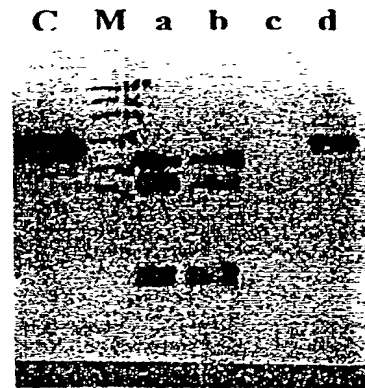


FIG. 9b

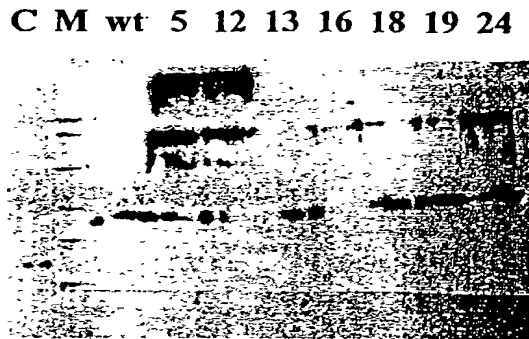


Fig. 10a

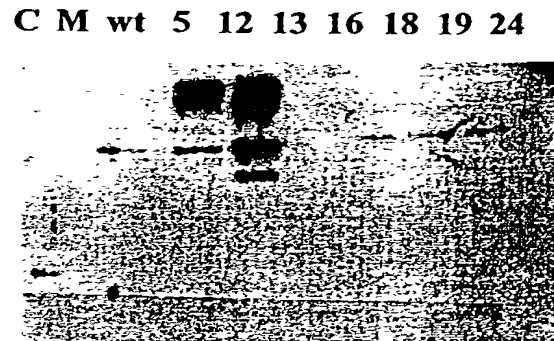


Fig. 10b

C M wt 2 5 7 9 10 11

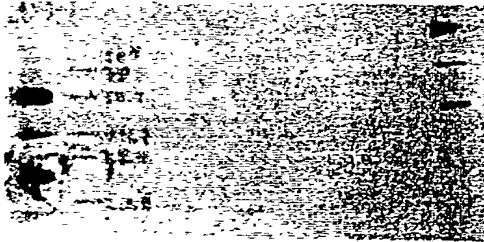


FIG. 11a

C M wt 13 14 15 17 18 19



FIG. 11c



FIG. 11b

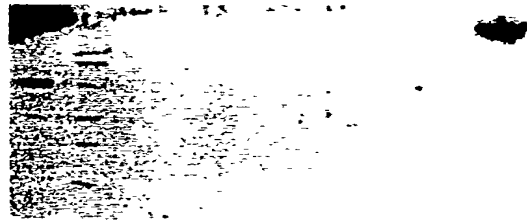


FIG. 11d

I hereby claim the benefit under 35 U.S.C. Section 119(e) of any United States provisional application(s) listed below:

(Application Serial No.)

(Filing Date)

(Application Serial No.)

(Filing Date)

(Application Serial No.)

(Filing Date)

I hereby claim the benefit under 35 U.S.C. Section 120 of any United States application(s), or Section 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. Section 112. I acknowledge the duty to disclose to the United States Patent and Trademark Office all the information known to me to be material to patentability as defined in Title 37, C.F.R., Section 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application:

PCT/IL00/00330

7 JUNE 2000

(Application Serial No.)

(Filing Date)

(Status)
(patented, pending, abandoned)

09/329,234

10 JUNE 1999

(Application Serial No.)

(Filing Date)

PENDING

(Status)
(patented, pending, abandoned)

(Application Serial No.)

(Filing Date)

(Status)
(patented, pending, abandoned)

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith. *(list name and registration number)*

SOL SHEINBEIN
Registration Number 25,457

MARTIN D. MOYNIHAN
Registration Number 40,338

Send Correspondence to: **G. E. EHRLICH (1995) LTD.**
c/o ANTHONY CASTORINA
SUITE 207
2001 JEFFERSON DAVIS HIGHWAY
ARLINGTON, VIRGINIA 22202

Direct Telephone Calls to: *(name and telephone number)*

Anthony Castorina Tel. No. **703-4151581**
Fax No. **703-4154864**

FULL NAME OF SOLE OR FIRST INVENTOR	<u>Ziv SHANI</u>
Sole or first inventor's signature	<u>[Signature]</u> Date <u>27 November 2001</u>
Residence	: 4 Paldi Street, 76248 Rehovot, Israel <u>ILX</u>
Citizenship	: ISRAELI
Post Office Address	: 4 Paldi Street, 76248 Rehovot, Israel

FULL NAME OF SECOND INVENTOR, IF ANY	<u>Oded SHOSEYOV</u>
Second inventor's signature	<u>[Signature]</u> Date <u>27 November 2001</u>
Residence	: 5 Haerez Street, 72910 Karme Yosef, Israel <u>ILX</u>
Citizenship	: ISRAELI
Post Office Address	: 5 Haerez Street, 72910 Karme Yosef, Israel

SEQUENCE LISTING

<110> Ziv Shani *et al.*

<120> PROCESS OF EXPRESSING AND ISOLATING RECOMBINANT PROTEINS AND RECOMBINANT
PROTEIN PRODUCTS FROM PLANTS, PLANT DERIVED TISSUES OR CULTURED PLANT CELLS

<130> 20101

<150> 09/329,234-

<151> June 10, 1999

<160> 26

<170> PatIn Ver. 2.1

<210> 1

<211> 4

<212> PRT

<213> unknown

<400> 1

His Asp Glu Leu

4

<210> 1

<211> 4

<212> PRT

<213> unknown

<400> 1

Lys Asp Glu Leu

4

<210> 3

<211> 150

<212> DNA

<213> Arabidopsis thaliana

<400> 3

ATGGCGCGAA AATCCCTAAT TTTCCCGGTG ATTTTGCTCG CCGTTCTTCT CTTCTCTCCG 60

CCGATTTACT CCGCCGGTCA CGATTACCGC GACGCTCTCC GTAAAAGCAT TCTCTTCTTC 120

GAAGGTCAAC GTTCCGGTAA ACTCCCTCCA 150

<210> 4

<211> 29

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial Sequence:PCR primer

<400> 4

AAAAAGATCT ATGGCGCGAA AATCACTAA 29

<210> 5

<211> 29

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial Sequence:PCR primer

<400> 5

AAAATCTAGA TTTACGGAGA GCGTCGCG 29

<210> 6

<211> 1305

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial Sequence:construct

<400> 6

CCATGGCGGC GGTAGAAAAT AAAGAAGAAA CACCAGAAAC ACCAGAACT GATTCAGAAG 60
AAGAAGTAAC AATCAAAGCT AACCTAATCT TTGCAAATGG AAGCACACAA ACTGCAGAAT 120
TCAAAGGAAC ATTTGAAAAA GCAACATCAG AAGCTTATGA GTATGCAGAT ACTTTGAAGA 180
AAGACAATGG AGAATATACT GTAGATGTTG CAGATAAAGG TTATACTTTA AATATTAAAT 240
TTGCTGGAAG AGAAAAACA CCAGAAGAAC CAAAAGAAGA AGTTACTATT AAAGCAAAC 300
TAATCTATGC AGATGGAAAA ACACAAACAG CAGAATTCAA AGGAACATTT GAAGAAGCAA 360
CAGCAGAAGC ATACAGATAT GCAGATGCAT TAAAGAAGGA CAATGGAGAA TATACAGTAG 420
ACGTTGCAGA TAAAGGTTAT ACTTTAAATA TTAAATTGTC TGGAAAAGAA AAAACACCAG 480
AAGAACCAA AGAAGAAGTT ACTATTAAAG CAACTTAAT CTATGCAGAT GGAAAAACAC 540
AAACAGCAGA ATTCAAAGGA ACATTGAAG AAGCAACAGC AGAAGCATAC AGATATGCTG 600
ACTTATTAGC AGCAAAAGAA AATGGTAAAT ATACAGTAGA CGTTGCAGAT AAAGGTTATA 660
CTTTAAATAT TAAATTTGCT GGAAAAGAAA AAACACCAGA AGAACCAGAA GAAGAAGTTA 720
CTATTAAAGC AACTTAATC TATGCAGATG GAAAACTCA AACAGCAGAG TTCAAAGGAA 780
CATTTGCAGA AGCAACAGCA GAAGCATACA GATACGCTGA CTTATTAGCA AAAGAAAATG 840
GTAAATATAC AGCAGACTTA GAAGATGGTG GATACACTAT TAATATTAGA TTTGCAGGTA 900
AGAAAGTTGA CGAAAAACCA GAAGGGATCC CTCCGACGCC GACCCCGACT AGTGGTCCGG 960
CCGGGTGCCA GGTGCTGTGG GCGTCAACC AGTGGAACAC CGGCTTCACC GCGAACGTCA 1020

<220>

<223> Description of Artificial Sequence:PCR primer

<400> 10

AAAAGTCGAC TTAAAGTTCA TCATGTCCAA CGGTGCAAGG GGC 43

<210> 11

<211> 1418

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial Sequence:construct

<400> 11

GGGGATCTAT GGC GCGAAAA TCCCTAATTT TCCCGGTGAT TTGCTCGCC GTTCTTCTCT 60
TCTCTCCGCC GATTTACTCC GCCGGTCACG ATTACCGCGA CGCTCTCCGT AAATCTAGAA 120
TGGCGGCGGT AGAAAATAAA GAAGAAACAC CAGAAACACC AGAACTGAT TCAGAAGAAG 180
AAGTAACAAT CAAAGCTAAC CTAATCTTG CAAATGGAAG CACACAACT GCAGAATTCA 240
AAGGAACATT TGA AAAAGCA ACATCAGAAG CTTATGAGTA TGCAGATACT TTGAAGAAAG 300
ACAATGGAGA ATATACTGTA GATGTTGCAG ATAAAGGTTA TACTTTAAAT ATTAAATTTG 360
CTGGA AAAAGA AAAACACCA GAAGAACCA AAGAAGAAGT TACTATTAAA GCAAACCTTA 420
TCTATGCAGA TGGAAAACA CAAACAGCAG AATTC AAAGG AACATTTGAA GAAGCAACAG 480
CAGAAGCATA CAGATATGCA GATGCATTAA AGAAGGACAA TGGAGAATAT ACAGTAGACG 540
TTGCAGATAA AGGTTATACT TTAAATATTA AATTTGCTGG AAAAGAAAAA ACACCAGAAG 600
AACC AAAAGA AGAAGTTACT ATTAAGCAA ACTTAATCTA TGCAGATGGA AAAACACAAA 660
CAGCAGAATT CAAAGGAACA TTTGAAGAAG CAACAGCAGA AGCATACAGA TATGCTGACT 720
TATTAGCAGC AAAAGAAAAT GGTAAATATA CAGTAGACGT TGCAGATAAA GGT TATACTT 780
TAAATATTAA ATTTGCTGGA AAAGAAAAA CACCAGAAGA ACCAAAAGAA GAAGTTACTA 840
TTAAAGCAAA CTTAATCTAT GCAGATGGAA AACTCAAAC AGCAGAGTTC AAAGGAACAT 900
TTGCAGAAGC AACAGCAGAA GCATACAGAT ACGCTGACTT ATTAGCAAAA GAAAATGGTA 960
AATATACAGC AGACTTAGAA GATGGTGGAT AACTATTAA TATTAGATTT GCAGGTAAGA 1020
AAGTTGACGA AAAACCAGAA GGGATCCCTC CGACGCCGAC CCCGACTAGT GCTAGCGGTC 1080
CAGCCGGCTG CCAGGTTCTG TGGGTGTTA ACCAGTGGAA CACCGGTTTC ACCGCTCAGG 1140
TTACCGTTAA AAACACGGGC TCAGCTCCGG TTGACGGTTG GACCCTGACC TTCTCTTTTC 1200
CCTCGGGTCA GCAGGTA ACT CAGGCTTGGT CATCTACAGT TACCCAGTCT GGATCCGCTG 1260
TTACAGTTCTG TAACGCTCCG TGGAAACGTA ATATTCCTGC AGGTGGAACC GCTCAGTTCTG 1320
GTTTCCAAGG TTCTCACACC GGTACCAACG CGGCGCCAAC CGCTTTCTCT CTGAACGGGG 1380
CCCCTTGAC CGTTGGACAT GATGAAC TTT AAGTCGAC 1418

<210> 12

<211> 467

Gln Trp Asn Thr Gly Phe Thr Ala Gln Val Thr Val Lys Asn Thr Gly

WO 00/77174

PCT/IL00/00330

6

370	375	380
Ser Ala Pro Val Asp Gly Trp Thr Leu Thr Phe Ser Phe Pro Ser Gly		
385	390	395 400
Gln Gln Val Thr Gln Ala Trp Ser Ser Thr Val Thr Gln Ser Gly Ser		
	405	410 415
Ala Val Thr Val Arg Asn Ala Pro Trp Asn Gly Asn Ile Pro Ala Gly		
	420	425 430
Gly Thr Ala Gln Phe Gly Phe Gln Gly Ser His Thr Gly Thr Asn Ala		
	435	440 445
Ala Pro Thr Ala Phe Ser Leu Asn Gly Ala Pro Cys Thr Val Gly His		
	450	455 460
Asp Glu Leu		
465		

<210> 13
<211> 25
<212> DNA
<213> Artificial sequence
<220>
<223> Description of Artificial Sequence:PCR primer
<400> 13
CTAGTCATGA TGAAC TTAA GAGCT 25

<210> 14
<211> 17
<212> DNA
<213> artificial sequence
<220>
<223> Description of Artificial Sequence:PCR primer
<400> 14
CTTAAAGTTC ATCATGA 17

<210> 15
<211> 24
<212> DNA
<213> Artificial sequence
<220>
<223> Description of Artificial Sequence:PCR primer
<400> 24
AAAACCCGGG ATGGCGCGAA AATC 24

WO 00/77174

PCT/IL00/00330

7

<210> 16

<211> 33

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial Sequence:PCR primer

<400> 16

AAAAGACGTC TTACGGAGAG CGTCGCGGTA ATC 33

<210> 17

<211> 1560

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial Sequence:construct

<400> 17

GGCCCCCTC GAGCCCGGA TGGCGCGAAA ATCCCTAATT TTCCCGGTGA TTTGCTCGC 60
CGTTCTTCTC TTCTCTCCGC CGATTACTC CGCCGGTCAC GATTACCGCG ACGCTCTCCG 120
TAAGACGTCC GGTCCAGCCG GCTGCCAGGT TCTGTGGGGT GTTAACCAGT GGAACACCGG 180
TTTCACCGCT CAGGTTACCG TTA AAAACAC GGGCTCAGCT CCGGTTGACG GTTGGACCCT 240
GACCTTCTCT TTTCCCTCGG GTCAGCAGGT AACTCAGGCT TGGTCATCTA CAGTTACCCA 300
GTCTGGATCC GCTGTTACAG TTCGTAACGC TCCGTGGAAC GGTAATATTC CTGCAGGTGG 360
AACCGCTCAG TTCGGTTTCC AAGGTTCTCA CACCGGTACC AACCGGGCGC CAACCGCTTT 420
CTCTCTGAAC GGGGCCCCCT GCACCGTTGG TCCGACTACC TCACCTACAA CGCGTAAGCT 480
CTGCAGCCTG GACAACGGGG ACTGTGACCA GTTCTGCCAC GAGGAACAGA ACTCTGTGGT 540
GTGCTCTTGC GCCC CGGGGT ACACCCTGGC TGACAACGGC AAGGCCTGCA TTCCACAGG 600
GCCCTACCCC TGTGGGAAAC AGACCCTGGA ACGCAGGAAG AGGTCAGTGG CCCAGGCCAC 660
CAGCAGCAGC GGGGAGGCCC CTGACAGCAT CACATGGAAG CCATATGATG CAGCCGACCT 720
GGACCCACC GAGAACCCCT TCGACCTGCT TGACTTTGAT CAGACGCAGC CTGAGAGGGG 780
CGACAACAAC ATTGAAGGTC GTATCGTGGG AGGCCAGGAA TGCAAGGACG GGGAGTGTCC 840
CTGGCAGGCC CTGCTCATCA ATGAGGAAAA CGAGGGTTTC TGTGGTGGAA CCATTCTGAG 900
CGAGTTCTAC ATCCTAACGG CAGCCCACTG TCTCTACCAA GCCAAGAGAT TCAAGGTGAG 960
GGTAGGGGAC CGGAACACGG AGCAGGAGGA GGGCGGTGAG GCGGTGCACG AGGTGGAGGT 1020
GGTCATCAAG CACAACCGGT TCACAAAGGA GACCTATGAC TTCGACATCG CCGTGCTCCG 1080
GCTCAAGACC CCCATCACCT TCCGCATGAA CGTGGCGCCT GCCTGCCTCC CCGAGCGTGA 1140
CTGGGCCGAG TCCACGCTGA TGACGCAGAA GACGGGGATT GTGAGCGGCT TCGGGCGCAC 1200
CCACGAGAAG GGCCGGCAGT CCACCAGGCT CAAGATGCTG GAGGTGCCCT ACGTGGACCG 1260

Met 1	Ala 2	Arg 3	Lys 4	Ser 5	Leu 6	Ile 7	Phe 8	Pro 9	Val 10	Ile 11	Leu 12	Leu 13	Ala 14	Val 15	Leu 16
Leu 17	Phe 18	Ser 19	Pro 20	Pro 21	Ile 22	Tyr 23	Ser 24	Ala 25	Gly 26	His 27	Asp 28	Tyr 29	Arg 30	Asp 31	Ala 32
Leu 33	Arg 34	Lys 35	Thr 36	Ser 37	Gly 38	Pro 39	Ala 40	Gly 41	Cys 42	Gln 43	Val 44	Leu 45	Trp 46	Gly 47	Val 48
Asn 49	Gln 50	Trp 51	Asn 52	Thr 53	Gly 54	Phe 55	Thr 56	Ala 57	Gln 58	Val 59	Thr 60	Val 61	Lys 62	Asn 63	Thr 64
Gly 65	Ser 66	Ala 67	Pro 68	Val 69	Asp 70	Gly 71	Trp 72	Thr 73	Leu 74	Thr 75	Phe 76	Ser 77	Phe 78	Pro 79	Ser 80
Gly 81	Gln 82	Gln 83	Val 84	Thr 85	Gln 86	Ala 87	Trp 88	Ser 89	Ser 90	Thr 91	Val 92	Thr 93	Gln 94	Ser 95	Gly 96
Ser 97	Ala 98	Val 99	Thr 100	Val 101	Arg 102	Asn 103	Ala 104	Pro 105	Trp 106	Asn 107	Gly 108	Asn 109	Ile 110	Pro 111	Ala 112
Gly 113	Gly 114	Thr 115	Ala 116	Gln 117	Phe 118	Gly 119	Phe 120	Gln 121	Gly 122	Ser 123	His 124	Thr 125	Gly 126	Thr 127	Asn 128
Ala 129	Ala 130	Pro 131	Thr 132	Ala 133	Phe 134	Ser 135	Leu 136	Asn 137	Gly 138	Ala 139	Pro 140	Cys 141	Thr 142	Val 143	Gly 144
Pro 145	Thr 146	Thr 147	Ser 148	Pro 149	Thr 150	Thr 151	Arg 152	Lys 153	Leu 154	Cys 155	Ser 156	Leu 157	Asp 158	Asn 159	Gly 160
Asp 161	Cys 162	Asp 163	Gln 164	Phe 165	Cys 166	His 167	Glu 168	Glu 169	Gln 170	Asn 171	Ser 172	Val 173	Val 174	Cys 175	Ser 176
Cys 177	Ala 178	Arg 179	Gly 180	Tyr 181	Thr 182	Leu 183	Ala 184	Asp 185	Asn 186	Gly 187	Lys 188	Ala 189	Cys 190	Ile 191	Pro 192
Thr 193	Gly 194	Pro 195	Tyr 196	Pro 197	Cys 198	Gly 199	Lys 200	Gln 201	Thr 202	Leu 203	Glu 204	Arg 205	Arg 206	Lys 207	Arg 208
Ser 209	Val 210	Ala 211	Gln 212	Ala 213	Thr 214	Ser 215	Ser 216	Ser 217	Gly 218	Glu 219	Ala 220	Pro 221	Asp 222	Ser 223	Ile 224
Thr 225	Trp 226	Lys 227	Pro 228	Tyr 229	Asp 230	Ala 231	Ala 232	Asp 233	Leu 234	Asp 235	Pro 236	Thr 237	Glu 238	Asn 239	Pro 240
Phe 241	Asp 242	Leu 243	Leu 244	Asp 245	Phe 246	Asp 247	Gln 248	Thr 249	Gln 250	Pro 251	Glu 252	Arg 253	Gly 254	Asp 255	Asn 256
Asn 257	Ile 258	Glu 259	Gly 260	Arg 261	Ile 262	Val 263	Gly 264	Gly 265	Gln 266	Glu 267	Cys 268	Lys 269	Asp 270	Gly 271	Glu 272

9

Cys Pro Trp Gln Ala Leu Leu Ile Asn Glu Glu Asn Glu Gly Phe Cys
 275 280 285
 Gly Gly Thr Ile Leu Ser Glu Phe Tyr Ile Leu Thr Ala Ala His Cys
 290 295 300
 Leu Tyr Gln Ala Lys Arg Phe Lys Val Arg Val Gly Asp Arg Asn Thr
 305 310 315 320
 Glu Gln Glu Glu Gly Gly Glu Ala Val His Glu Val Glu Val Val Ile
 325 330 335
 Lys His Asn Arg Phe Thr Lys Glu Thr Tyr Asp Phe Asp Ile Ala Val
 340 345 350
 Leu Arg Leu Lys Thr Pro Ile Thr Phe Arg Met Asn Val Ala Pro Ala
 355 360 365
 Cys Leu Pro Glu Arg Asp Trp Ala Glu Ser Thr Leu Met Thr Gln Lys
 370 375 380
 Thr Gly Ile Val Ser Gly Phe Gly Arg Thr His Glu Lys Gly Arg Gln
 385 390 395 400
 Ser Thr Arg Leu Lys Met Leu Glu Val Pro Tyr Val Asp Arg Asn Ser
 405 410 415
 Cys Lys Leu Ser Ser Ser Phe Ile Ile Thr Gln Asn Met Phe Cys Ala
 420 425 430
 Gly Tyr Asp Thr Lys Gln Glu Asp Ala Cys Gln Gly Asp Ser Gly Gly
 435 440 445
 Pro His Val Thr Arg Phe Lys Asp Thr Tyr Phe Val Thr Gly Ile Val
 450 455 460
 Ser Trp Gly Glu Gly Cys Ala Arg Lys Gly Lys Tyr Gly Ile Tyr Thr
 465 470 475 480
 Lys Val Thr Ala Phe Leu Lys Trp Ile Asp Arg Ser Met Lys Thr Arg
 485 490 495
 Gly Leu Pro Lys Ala Lys Pro Thr Ser His Asp Glu Leu
 500 505

<210> 19

<211> 32

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial Sequence:PCR primer

<400> 19

AAAAGACGTC GGCTAGCGGA ATAATGGTAG CG 32

<210> 20

<211> 27

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial Sequence:PCR primer

<400> 20

AAAAACGCGT TGGGGATGGG GTCGGAC 27

<210> 21

<211> 1860

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial Sequence:construct

<400> 21

GGCCCCCTC GAGCCCGGA TGGCGGAAA ATCCCTAATT TTCCCGGTGA TTTTGCTCGC 60
CGTTCTTCTC TTCTCTCCGC CGATTTACTC CGCCGGTCAC GATTACCGCG ACGCTCTCCG 120
TAAGACGTCG GCTAGCGGAA TAATGGTAGC GACAGCAAAA TACGGAACAC CGGTCATCGA 180
TGGAGAGATA GACGAGATCT GGAACACGAC AGAGGAGATA GAGACGAAAG CGGTGGCCAT 240
GGGATCGCTT GACAAGAAG CGACAGCGAA AGTGAGGGTG CTGTGGGACG AGAACTACCT 300
GTACGTACTT GCAATCGTGA AAGACCCCGT TCTGAACAAA GACAACAGCA ACCCGTGGGA 360
ACAGGATTCC GTGGAGATCT TCATCGACGA GAACAACCAC AAGACAGGAT ACTACGAAGA 420
CGACGACGCA CAGTTCAGGG TGAATACAT GAACGAGCAG ACGTTTGGA CGGGAGGAAG 480
TCCAGCGAGG TTCAAGACAG CGGTGAAACT GATCGAAGGA GGATACATAG TTGAGGCAGC 540
GATCAAGTGG AAGACGATCA AACCCACACC GAACACGGTG ATAGGATTCA ACATCCAGGT 600
GAACGATGCG AACGAGAAAG GGCAGAGGGT CGGTATCATC TCCTGGAGCG ATCCCACAAA 660
CAACAGCTGG AGAGATCCTT CAAAGTTCGG TAACCTCAGG CTCATCAAGG GATCTGGTCC 720
GACCCCATCC CCAACGCGTA AGCTCTGCAG CCTGGACAAC GGGGACTGTG ACCAGTTCTG 780
CCACGAGGAA CAGAACTCTG TGGTGTGCTC CTGCGCCCGC GGGTACACCC TGGCTGACAA 840
CGGCAAGGCC TGCATTCCCA CAGGGCCCTA CCCCTGTGGG AAACAGACCC TGGAACGCAG 900
GAAGAGGTCA GTGGCCCAGG CCACCAGCAG CAGCGGGGAG GCCCCTGACA GCATCACATG 960
GAAGCCATAT GATGCAGCCG ACCTGGACCC CACCGAGAAC CCCTTCGACC TGCTTGACTT 1020
TGATCAGACG CAGCCTGAGA GGGGCGACAA CAACATTGAA GGTCGTATCG TGGGAGGCCA 1080
GGAATGCAAG GACGGGGAGT GTCCCTGGCA GGCCCTGCTC ATCAATGAGG AAAACGAGGG 1140
TTTCTGTGGT GGAACCATTC TGAGCGAGTT CTACATCCTA ACGGCAGCCC ACTGTCTCTA 1200
CCAAGCCAAG AGATTCAAGG TGAGGGTAGG GGACCGGAAC ACGGAGCAGG AGGAGGGCGG 1260
TGAGGCGGTG CACGAGGTGG AGGTGGTCAT CAAGCACAAC CGGTTACAAA AGGAGACCTA 1320
TGACTTCGAC ATCGCCGTGC TCCGGCTCAA GACCCCCATC ACCTTCGCA TGAACGTGGC 1380
GCCTGCCTGC CTCCCCGAGC GTGACTGGGC CGAGTCCACG CTGATGACGC AGAAGACGGG 1440
GATTGTGAGC GGCTTCGGGC GCACCCACGA GAAGGGCCGG CAGTCCACCA GGCTCAAGAT 1500
GCTGGAGGTG CCCTACGTGG ACCGCAACAG CTGCAAGCTG TCCAGCAGCT TCATCATCAC 1560
CCAGAACATG TTCTGTGCCG GCTACGACAC CAAGCAGGAG GATGCCTGCC AGGGGGACAG 1620
CGGGGGCCCG CACGTCACCC GCTTCAAGGA CACCTACTTC GTGACAGGCA TCGTCAGCTG 1680

Met 1	Ala 2	Arg 3	Lys 4	Ser 5	Leu 6	Ile 7	Phe 8	Pro 9	Val 10	Ile 11	Leu 12	Leu 13	Ala 14	Val 15	Leu 16
Leu 17	Phe 18	Ser 19	Pro 20	Pro 21	Ile 22	Tyr 23	Ser 24	Ala 25	Gly 26	His 27	Asp 28	Tyr 29	Arg 30	Asp 31	Ala 32
Leu 33	Arg 34	Lys 35	Thr 36	Ser 37	Ala 38	Ser 39	Gly 40	Ile 41	Met 42	Val 43	Ala 44	Thr 45	Ala 46	Lys 47	Tyr 48
Gly 49	Thr 50	Pro 51	Val 52	Ile 53	Asp 54	Gly 55	Glu 56	Ile 57	Asp 58	Glu 59	Ile 60	Trp 61	Asn 62	Thr 63	Thr 64
Glu 65	Glu 66	Ile 67	Glu 68	Thr 69	Lys 70	Ala 71	Val 72	Ala 73	Met 74	Gly 75	Ser 76	Leu 77	Asp 78	Lys 79	Asn 80
Ala 81	Thr 82	Ala 83	Lys 84	Val 85	Arg 86	Val 87	Leu 88	Trp 89	Asp 90	Glu 91	Asn 92	Tyr 93	Leu 94	Tyr 95	Val 96
Leu 97	Ala 98	Ile 99	Val 100	Lys 101	Asp 102	Pro 103	Val 104	Leu 105	Asn 106	Lys 107	Asp 108	Asn 109	Ser 110	Asn 111	Pro 112
Trp 113	Glu 114	Gln 115	Asp 116	Ser 117	Val 118	Glu 119	Ile 120	Phe 121	Ile 122	Asp 123	Glu 124	Asn 125	Asn 126	His 127	Lys 128
Thr 129	Gly 130	Tyr 131	Tyr 132	Glu 133	Asp 134	Asp 135	Ala 136	Gln 137	Phe 138	Val 139	Val 140	Asn 141	Tyr 142	Met 143	
Asn 144	Glu 145	Gln 146	Thr 147	Phe 148	Gly 149	Thr 150	Gly 151	Gly 152	Ser 153	Pro 154	Ala 155	Arg 156	Phe 157	Lys 158	Thr 159
Ala 160	Val 161	Lys 162	Leu 163	Ile 164	Glu 165	Gly 166	Gly 167	Tyr 168	Ile 169	Val 170	Glu 171	Ala 172	Ala 173	Ile 174	Lys 175
Trp 176	Lys 177	Thr 178	Ile 179	Lys 180	Pro 181	Thr 182	Pro 183	Asn 184	Thr 185	Val 186	Ile 187	Gly 188	Phe 189	Asn 190	Ile 191
Gln 192	Val 193	Asn 194	Asp 195	Ala 196	Asn 197	Glu 198	Lys 199	Gly 200	Gln 201	Arg 202	Val 203	Gly 204	Ile 205	Ile 206	Ser 207
Trp 208	Ser 209	Asp 210	Pro 211	Thr 212	Asn 213	Asn 214	Ser 215	Trp 216	Arg 217	Asp 218	Pro 219	Ser 220	Lys 221	Phe 222	Gly 223
Asn 224	Leu 225	Arg 226	Leu 227	Ile 228	Lys 229	Gly 230	Ser 231	Gly 232	Pro 233	Thr 234	Pro 235	Ser 236	Pro 237	Thr 238	Arg 239
Lys 240	Leu 241	Cys 242	Ser 243	Leu 244	Asp 245	Asn 246	Gly 247	Asp 248	Cys 249	Asp 250	Gln 251	Phe 252	Cys 253	His 254	Glu 255
Glu 256	Gln 257	Asn 258	Ser 259	Val 260	Val 261	Cys 262	Ser 263	Cys 264	Ala 265	Arg 266	Gly 267	Tyr 268	Thr 269	Leu 270	Ala 271
Asp 272	Asn 273	Gly 274	Lys 275	Ala 276	Cys 277	Ile 278	Pro 279	Thr 280	Gly 281	Pro 282	Tyr 283	Pro 284	Cys 285	Gly 286	Lys 287
Gln 288	Thr 289	Leu 290	Glu 291	Arg 292	Arg 293	Lys 294	Arg 295	Ser 296	Val 297	Ala 298	Gln 299	Ala 300	Thr 301	Ser 302	Ser 303

12

290	295	300
Ser Gly Glu Ala Pro Asp	Ser Ile Thr Trp Lys Pro Tyr Asp Ala Ala	
305	310	315 320
Asp Leu Asp Pro Thr Glu Asn Pro Phe Asp Leu Leu Asp Phe Asp Gln		
	325	330 335
Thr Gln Pro Glu Arg Gly Asp Asn Asn Ile Glu Gly Arg Ile Val Gly		
	340	345 350
Gly Gln Glu Cys Lys Asp Gly Glu Cys Pro Trp Gln Ala Leu Leu Ile		
	355	360 365
Asn Glu Glu Asn Glu Gly Phe Cys Gly Gly Thr Ile Leu Ser Glu Phe		
	370	375 380
Tyr Ile Leu Thr Ala Ala His Cys Leu Tyr Gln Ala Lys Arg Phe Lys		
	385	390 395 400
Val Arg Val Gly Asp Arg Asn Thr Glu Gln Glu Glu Gly Gly Glu Ala		
	405	410 415
Val His Glu Val Glu Val Val Ile Lys His Asn Arg Phe Thr Lys Glu		
	420	425 430
Thr Tyr Asp Phe Asp Ile Ala Val Leu Arg Leu Lys Thr Pro Ile Thr		
	435	440 445
Phe Arg Met Asn Val Ala Pro Ala Cys Leu Pro Glu Arg Asp Trp Ala		
	450	455 460
Glu Ser Thr Leu Met Thr Gln Lys Thr Gly Ile Val Ser Gly Phe Gly		
	465	470 475 480
Arg Thr His Glu Lys Gly Arg Gln Ser Thr Arg Leu Lys Met Leu Glu		
	485	490 495
Val Pro Tyr Val Asp Arg Asn Ser Cys Lys Leu Ser Ser Ser Phe Ile		
	500	505 510
Ile Thr Gln Asn Met Phe Cys Ala Gly Tyr Asp Thr Lys Gln Glu Asp		
	515	520 525
Ala Cys Gln Gly Asp Ser Gly Gly Pro His Val Thr Arg Phe Lys Asp		
	530	535 540
Thr Tyr Phe Val Thr Gly Ile Val Ser Trp Gly Glu Gly Cys Ala Arg		
	545	550 555 560
Lys Gly Lys Tyr Gly Ile Tyr Thr Lys Val Thr Ala Phe Leu Lys Trp		
	565	570 575
Ile Asp Arg Ser Met Lys Thr Arg Gly Leu Pro Lys Ala Lys Pro Thr		
	580	585 590
Ser His Asp Glu Leu		
	595	

<210> 23

<211> 28

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial Sequence:PCR primer

<400> 23

AAAACCATGG CGGCGGTAGA AAATAAAG 28

WO 00/77174

PCT/IL00/00330

13

<210> 24

<211> 29

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial Sequence:PCR primer

<400> 24

AAAAGGATCC CTTCTGGTTT TTCGTCAAC 29

<210> 25

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> Description of Artificial Sequence:PCR primer

<400> 25

AAAACCCGGG ATGGCGCGAA AATC 24

<210> 26

<211> 21

<212> DNA

<213> Artificial sequence

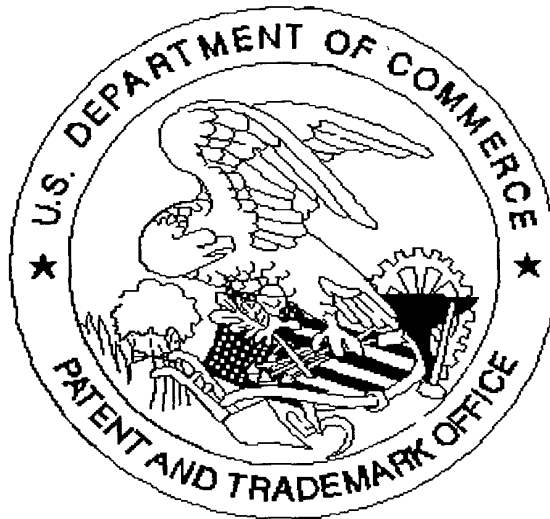
<220>

<223> Description of Artificial Sequence:PCR primer

<400> 26

TGC GTTCCAG GGTCTGTTTC C 21

United States Patent & Trademark Office
Office of Initial Patent Examination -- Scanning Division



Application deficiencies found during scanning:

☐ Page(s) _____ of _____ were not present
for scanning. (Document title)

☐ Page(s) _____ of _____ were not present
for scanning. (Document title)

*Scanned copy is best available. Drawings fig. 5-8b
are very dark.*